| | CATTCCGATT | ATCATTTCAT | ATAAAGAAGG | TTTACATATT | ATTAAAGATT | TAATTGTTGC | 15900 |
|------------|------------------|--------------|------------|--------------|------------|------------|-------|
| | GACATTACGA | GCAGTTGTGC | AATTAATCAT | TTTGGGATTT | TTGCTGCATT | ATATTTTTAA | 15960 |
| 5 | AATAAACGAT | AAATGGCTGC | TTATTTTATG | TGTATTGGTC | ATTATTATTA | ATGCATCATG | 16020 |
| | GAATACAATT | AGTCGAGCAT | CACCAGTGAT | GCATCATGTG | TTTTGGATAT | CATTTCTAGC | 16080 |
| 10 | TATCTTCATT | GGAACGGCAT | TACCGCTTGC | AGGTACTATT | GCGACAGGGG | CCATTCAATT | 16140 |
| 70 | TACCGCAAAT | GAAGTTATAC | CTATCGGCGG | CATGCTTGCA | AATAATGGCT | TGATTGCAAT | 16200 |
| | TAATTTAGCT | TACCAGAATT | TAGATCGTGC | ATTCGTACAA | GATGGTACTA | ATATTGAATC | 16260 |
| 15 | TAAATTATCA | CTTGCAGCTA | CACCTAAATT | GGCTTCTAAA | GGTGCAATAC | GTGAAAGTAT | 16320 |
| | TCGTTTAGCT | ATAGTGCCAA | CTATTGATTC | GGTTAAAACA | TATGGGCTTG | TGTCGATTCC | 16380 |
| | TGGTATGATG | ACAGGCTTAA | TTATTGGTGG | CGTACCACCT | TTACAAGCGA | TTAAATTTCA | 16440 |
| 20 | ATTGTTAGTC | GTGTTTATTC | ATACAACTGC | GACCATTATG | TCTGCTTTGA | TTGCGACATA | 16500 |
| | TTTAAGCTAT | GGTCAATTTT | TCAATGCAAG | ACATCAATTA | GTAGCACGAA | ATACTGATGT | 16560 |
| | TAAGAGTGAA | TCATGATAGA | TTTTACTGCA | TCAGATTTAG | GCATTAGTTT | TAATTGGAAA | 16620 |
| ?5 | TGAAGTGACG | CGCACATATA | GTATCGCTAT | TCATTAGCGC | AGCGAAAATA | TTCATAAAGG | 16680 |
| | CACGCATACT | TTGTAGTCAG | TTATCTGTTC | TGACATATAA | AGCGTGCGTG | CTTTTTTGGA | 16740 |
| | GTTATTGTTG | AAACTGAAGT | AATTATACAT | AATTATTAAA | TGACATACTT | GTGTTAATTT | 16800 |
| 30 | TTCAAATACT | GAAAAACAAT | TTCaATAATT | TTCCaATTAA | GCACAGAAAA | TTAAAGCAAA | 16860 |
| | ATATTATA | ATAGAACGGT | TATATATAA | nATTngTgCA | CACATTTTTT | AATAAATCGT | 16920 |
| 3 <i>5</i> | TATTCTAAGG | GAAATGAATA | TCGGAAATTT | TGTTTGAAAG | GAGTTTTAAA | TTGTCAATCA | 16980 |
| | TGCGACTATT | TACATTCATT | TTAAGTATTT | TTATCGTAGG | aatggttgaa | ATGATGGTTG | 17040 |
| | CAGGÃATTAT | GAACTTGATG | AGTCAGGACT | TACATGTATC | AGAAGCTGTC | GTTGGTCAAT | 17100 |
| 10 | TAGTGACAAT | GTACGCTTTA | ACATTTGCGA | TATGTGGACC | TATTCTGGTT | AAATTAACGA | 17160 |
| | ACCGTTTTTC | ATCAAGGCCT | GTATTATTAT | GGACATTACT | TATATTTATC | ATTGGTAATG | 17220 |
| | GCATTATTGC | TGTAGCGCCA | AATTTTTCaA | TATTAGTAGT | TGGTAGAATT | ATCTCATCTG | 17280 |
| 15 | CAGCAGCAGC | ACTAATTATC | GTAAAAGTAT | TAGCTATTAC | AGCGATGTTA | TCAGCACCTA | 17340 |
| | AAAATCGTGG | TAAAATGATT | GGACTTGTCT | ATACAGGGTT | TAGTGGTGCT | AATGTTTTTG | 17400 |
| | GTGTACCAAT | TGGAACGGTT | ATCGGCGATT | TAGTAGGTTG | GCGCTATACA | TTTCTATTCT | 17460 |
| 50 | TAATTATTGT | GAGTATTATT | GTTGGCTTCT | TGATGATGAT | CTATTTACCG | AAGGATCAGG | 17520 |
| | A A A TA CA A CC | 1.CCCCCCCCCC | | C1 CC1 TCTC1 | | | |

| | CAAACTCAGT | GACATTCGTC | TTTATAAATC | CACTTATTTT | ATCTAATGGT | CATGATATGT | 17700 |
|----|-------------|-------------|-------------|------------|------------|------------|-------|
| | CATTCGTTTC | ATTAGCACTT | CTAGTAAATG | GAATCGCTGG | CGTTATTGGA | ACATCATTAG | 17760 |
| 5 | GTGGTATATT | CTCCGATAAA | ATTACAAGTA | AGCGTTGGTT | AATGATTTCT | GTTTCTATTT | 17820 |
| | TTATCGTCAT | GATGTTACTT | ATGAATTTAA | TCTTACCTGG | TTCAGGTCTA | TTGTTAGCAG | 17880 |
| 10 | GACTATTTAT | TTGGAATATC | ATGCAATGGA | GTACTAATCC | AGCAGTGCAA | AGCGGTGTGA | 17940 |
| | TTCAACATGT | TGAAGGCGAC | ACAAGCCAAG | TAATGAGTTG | GAACATGTCT | AGTTTAAACG | 18000 |
| | CTGGTATTGG | TGTTGGAGGC | ATTATTGGAG | GCTTGGTCAT | GACACATGTT | TCTGTTCAAG | 18060 |
| 15 | CTATCACATA | TACGAGTGCC | ATCATTGGCG | CATTAGGATT | AATCGTTGTT | TTCACATTGA | 18120 |
| | AAAATAATCA | TTATGCTAAA | ACATTTAAAT | CATCATAATT | CTCATATGAm | AAGCACGCCT | 18180 |
| | GCTATCAAAT | TCAGGTGTGC | TTTTTTAGAT | GCGATAACGT | TATTGATATG | TGCGATAATA | 18240 |
| 20 | GCGACGTTCA | TTATGATACA | TCGGCCAAGG | CATTTTACCG | CTTTTAGCAA | AATTAGCTAA | 18300 |
| | ATCATTTTGC | ATTTGTCGAC | TTAAAAATTT | AAGGTGaGCA | GTTGTTGGaT | ATGAT | 18355 |
| | (2) INFORMA | TION FOR SE | O ID NO: 68 | 3 : | | • . | - |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1192 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

| 35 | CGCAAAGAAG | TACAAAAAAT | GTTTTTACAA | GAAGGTATTA | AAACACCTCA | ACCAATTATG | 60 |
|----|------------|------------|------------|------------|------------|------------|-----|
| 55 | ACTGCTTATA | ATCATAGTGA | AAACGgTGTT | TAGTAGTTTA | TAATACATGG | AGGTCATATT | 120 |
| | TAATGGCGTC | AAAATATGGA | ATAAATGATA | TAGTAGAAAT | GAAAAAACAA | CATGCGTGTG | 180 |
| 40 | GAACAAACCG | TTTTAAGATT | ATTAGAATGG | GTGCAGACAT | AAGAATTAAA | TGTGAAAATT | 240 |
| | GTCAAAGAAG | TATTATGATT | CCACGTCAAA | CGTTTGATAA | AAAACTTAAA | AAAATCATCG | 300 |
| | AATCTCATGA | TGATACACAA | AGATAGGAGA | ATGATTAATG | GCTTTAACAG | CAGGTATCGT | 360 |
| 45 | TGGATTGCCA | AACGTTGGTA | AATCAACATT | ATTTAATGCA | ATAACAAAAG | CAGGTGCTTT | 420 |
| | AGCAGCGAAC | TATCCATTCG | CTACGATTGA | TCCTAATGTA | GGGATAGTAG | AAGTGCCAGA | 480 |
| | TGCTAGATTA | CTTAAATTAG | AAGAAATGGT | TCAACCTAAA | AAGACATTGC | CGACTACATT | 540 |
| 50 | TGAATTTACA | GATATCGCTG | GTATTGTGAA | AGGTGCTTCA | AAGGGAGAAG | GGTTAGGTAA | 600 |
| | TAAATTCTTA | TCACATATTA | GAGAAGTAGA | TGCGATTTGT | CAGGTCGTTC | GTGCATTTGA | 660 |

| TAAT | ATGGAA | TTAGTACTAG | CGGACTTAGA | ATCTGTTGAG | AAACGTTTGC | CTAGAATTGA | 780 |
|-------|--------|------------|------------|------------|------------|------------|------|
| AAAA? | TAGCA | CGTCAAAAAG | ATAAGACTGC | TGAAATGGAA | GTACGTATTT | TAACAACTAT | 840 |
| TAAAC | SAAGCT | TTAGAAAATG | GTAAACCCGC | TCGTAGTATT | GACTTTAATG | AAGAAGATCA | 900 |
| CAAAA | rgggtg | AATCAAGCGC | AATTACTGAC | TTCTAAAAAA | ATGCTTTATA | TCGCTAATGT | 960 |
| TGGTC | BAAGAT | GAAATTGGTG | ATGATGATAA | TGATAAAGTA | AAAGCGATTC | GTGAATATGC | 1020 |
| AGCGC | CAAGAA | GACTCTGAAG | TGATTGTTAT | TAGTGCAAAA | ATTGAAGAAG | AAATTGCTAC | 1080 |
| ATTAC | SATGAT | GAAGATAAAG | AAATGTTCTT | AGAAGaTTTA | GGTATCGaAG | AACCAGGATT | 1140 |
| AGATO | grTTA | ATTAGGAMCA | cttatgaatt | ATTAGGNTTA | TCCACCATAA | TT | 1192 |
| | | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 69:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7494 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

| AATATAGCTG | CAATAGCATC | TCGTTTCATT | TGTATAATCA | ATTCCGGTTT | AAATATCAGT | 60 |
|------------|------------|------------|------------|------------|------------|------|
| GTGAACGTAA | GCACGACACA | GATTAAAAAT | AACACTGCCG | GAATGAGTCG | TTTCAATCGT | 120 |
| CGCTtCCAAA | ACTCTAGCAA | ATCGATTTTT | TGCGTCCGAT | AATACTCACT | TATCAACAAA | 180 |
| CTTGTTATTA | AATAACCTGA | AATAACGAAG | AATGTATCTA | CTCCTAAAAA | GCCCCCACTT | 240 |
| AACCATTGTG | CATTCAAGTG | ATAAATAATG | ATTCCTATAA | CTGCGAATGC | CCTCAATCCA | 300 |
| TCTAATCCAG | GTAAGTATCG | CGGGGAATAC | ATTTTTTCTA | AACGTTTAAA | GTCTTTTGTA | 360- |
| TCCATGTTAA | TAAACGCCCC | ATTTATTTTT | CTCTATTTTG | TAGTATATCA | CAATATTTTT | 420 |
| GAAAÄTAAAA | TATTGCACTG | aTTTTCATTA | ATTGATTTAA | CCCTTAATTA | AGATAGTTTT | 480 |
| AAATTTTTTA | TTAAGTAGAA | AACAATTATT | ACAGTTGATT | TCATTACTGC | AAACCACATA | 540 |
| TAAATTTGTC | GATTTTACTA | CATAACATAG | ATTATCATAG | ATTCTTGAAT | TTTTAGCAAA | 600 |
| ATAACTGTTA | TTTTCATTAT | ATTTTTACAA | AAAAAGGTTC | GTTTTATATT | TTATGCATCT | 660 |
| TACTGTAACA | GAATCATTAA | GATATGCTAT | TCGAATATAC | TTTTTCAAAA | TTTATATAAT | 720 |
| GAATAAATTA | ACATGTATTG | AAAAAAAAGC | GAAATGCAGC | CTATCCTCTA | ATGTAAACCA | 780 |
| AACGATATAT | CTCGTCAGAC | TTTATATTTA | AACGCTATGT | GTCACTTTTA | AAATGAATAT | 840 |
| TACTAAGATT | GTCATATCAA | TTATTATTGC | ATCGAATTAA | TCTTTTAAAT | TTCTGTAATA | 900 |

55

5

10

15

20

25

30

35

40

45

| | ACGGAAGTCA | TTATTAGAAT | AAAAATACTO | TGCACTAATA | AATTTATCA | TTGTTCCTAA | 1020 |
|----|---------------|-------------|--------------------|------------|-------------|-------------|------|
| | ATAAATACCA | TCGATATTTT | GTTCTTTACA | TGTCATTATA | ACTITATOTA | AAAGTTTTT | 1080 |
| 5 | ACCTATTTT | AAATTCCTAT | AACCTTTATO | AACAAACATT | TTTTTAAGTG | CAGACATATT | 1140 |
| | ATTATCTAGT | CTAATCAAAC | CTATAGTACC | AACAATATTT | TGaTGATTGT | TTATTGCAAG | 1200 |
| 10 | CCAAAATgCC (| CTCCATTATT | CAAATAGTTA | TGTTCGATGT | TCTCCAAATC | AGGTTGATCA | 1260 |
| ,, | TCTCTATCAA | ITTTTATAT | AATTCATTTT | TTTGAATCGA | TAAAATAAAC | TCGATTAGCT | 1320 |
| | CTTCCTTATA A | AGACCTATTA | TATTCAATTA | TGTTTATAGC | CATTTTTATC | TCCTTTTTCA | 1380 |
| 15 | TTTAATTTAA 1 | TTATAAAATG | TGCGTTTAGT | TTGTATCTAG | TGTACTCAGT | ACAGCCTCAA | 1440 |
| - | ATGAAGTTTC A | ATTCCACTTG | GCACTTAATA | AAGACAAGTA | TTTTAGCAGT | AATACAATAA | 1500 |
| | AGTCCAATAA A | TTTCCCTAA | CTTCAATATC | CACTTTTTAA | AAAATGTATT | TTTAATTAAT | 1560 |
| 20 | AAAAAAACTC T | CCCCAATTT | CTATGGGAAG | AGCTATATAT | TTAATGTCTA | AACATTACTT | 1620 |
| | TTATTTATTA T | GAAGGAATT | AGAATCCCCA | AGCACCȚAAA | CCTTGTGCTT | TGTATGCTTT | 1680 |
| | AACAGCTGCG T | TGATTTGTT | GGTCAACAGT | GTTTGTTGGA | CCCCAACCTG | GCATAGTTTG | 1740 |
| 25 | GAATAAACCT G | AAGCACCTG | ATGGGTTGTA | AGCATTTACT | TGACCATTTG | ATTCACGAGC | 1800 |
| | GATGATTGCA G | CCCATGTAG | AAGCTGAAAC | ACCAGTACGT | TGAGCCATGA | TTTGAGCTGC | 1860 |
| | TGATGAACCA G | TAGCACCTG | CAGTATTACC | ATTGCTTAAT | CTCACTGAAC | TTGAAGTAGT | 1920 |
| 30 | TGAAGTGCTG T | AGTTATGGT | AAGTTGGAGC | TGAAACAGCT | TCAACGTtTG | AGTTACTTGA | 1980 |
| | TTGTGCATTG T | AGCTTACTG | ATTGTACATT | TGAACCTTGG | TTGTATGAAG | TAGTGTAGTC | 2040 |
| 35 | TGCACCTGCA A | CGTTTGAGA . | AACCAGCAGT | TTGACCATTA | GCTGCTTCAT | AGCTCCATGA | 2100 |
| | CCATGTAGTA C | CATTTGAAG | TGAAGTTATA | TTGGAAACCA | TCTTTTACAA | AGTGGATGTC | 2160 |
| | ATATGCACCA T | CTTTGATTG | GAGCTGCATT | TAATTGATCT | TGGTGATTAT | GCGCTAAGTC | 2220 |
| 10 | AACTAAGTGT G | CTTGATCAA | CGTTTACTTC | AGCAGCGTGT | GCTTGATGTC | CTGTACCTGC | 2280 |
| | TGCGTAACCT G | TTACACCTA A | ATGCCACTGC | TAATGATGAT | GCCATAATTG | TCTTTTTCAT | 2340 |
| | AGTAAAAAAT C | CTCCAGTAA ' | TAATTGTnAG | TTTATGTTTT | TAGTAATTAT | AtTTTGaATT | 2400 |
| 15 | TGAATGTCGT A | GTgCAAGTT : | TAAATTGTCT | TTTATTTCTT | TCAACGGTAC | TCACTATATC | 2460 |
| | ACABAAAACC AG | GCCAGTAAA 1 | TACACTTTC | TTTACAAAAC | ATTACAATAT | CAAGTGTTAT | 2520 |
| | TTGLAATGTT G | AAATATGGC 1 | IGTTTTATA C | TGTAATGTGA | AATATGTGCC | CTTTAGAATC | 2580 |
| 50 | CAATCAACCC T | rgaaatagt (| CTTTAACACA | TAAGATTTTT | ACTATATTTA | GCTCAACTAT | 2640 |
| | TACAGCTTTC GT | CARTATTAC | AGATTGTATT | ттсттьсьт | ACCTCTA ATA | Throman and | |

| | TACACATGTA | TTGATTGCTA | TTATTGTTGT | ATATTCAAAG | TTTTAAAACA | CACATCTTTT | 2820 |
|----|------------|------------|------------|------------|------------|------------|------|
| | GTGAATTGTC | TTATCTTTTA | TTAGCGCAAA | TAAACTGCAG | CTCAATTATA | TTGTTCAACT | 2880 |
| 5 | TCATTCTCGC | AATTCACAAT | AACATTAAAT | AATTTTTGGT | CTCATATTTT | CAAAAAACAT | 2940 |
| | ACTGTTATTA | TCCCATGAAT | TTAAAAATAT | CATTAGTATA | TAAACGAAAC | ACTTTACGAT | 3000 |
| 10 | AAATGATATC | TGCAAGCCAA | GCTGTTACAA | ATGGTACAAC | AAAGAACGCT | ACTACAATTA | 3060 |
| 10 | GTAAGACACT | CAACCAAGCA | GAATCAACCT | CCATAAATTT | AAATGCATTA | ATCGGTCCTA | 3120 |
| | CCATTCCTAT | AAAACCAAAT | CCAGCTGACT | CTTTCGTTCC | ATGAATACCT | ACTAATGCTG | 3180 |
| 15 | ATACCAAACC | TGATACAATG | GCTGTCGTTA | ATATTGGTAA | CATAAGAATT | GGATATTTCA | 3240 |
| | CCATATTAGG | TATCATCATT | TTAACGCCTC | CAAAGAAGAC | GGATAACGGC | ACCCCTAAAC | 3300 |
| | GATTCACTTT | ACTTGTACCA | ATTATCAATA | CTGCTTCAGT | CGCGGAGATA | CCAATTGACG | 3360 |
| 20 | CTGATCCAGC | TGCTAAACCT | GTAATACCTA | TCGCAAAGGC | AATGGCCACA | GTTGATAGTG | 3420 |
| | GCGAAATAAT | AATAAGACTA | AATACCATTG | AAATCAAAAT | ACTCATGACA | ATCGGTTGTA | 3480 |
| | ATTCTGTAAA | ACCATTAACC | ATATTACCGA | TGGCTGTTGT | AATCATTTTC | GTATACGGCA | 3540 |
| ?5 | ATATTAAAAC | ACCAATTGCA | CCTGAAATAC | CGCCAACAAC | TGTTGGGAAT | ACAATCAATG | 3600 |
| | CCATACTACC | TACGCGATGT | TGAATAAGTA | AAATGAATAA | CACTGCAATC | GCTGCTGTAA | 3660 |
| | TCATTGTATT | AATTAAATCA | CCAATACCCG | TAATCATCCA | AGCACCATTT | TTAAACTGCG | 3720 |
| 30 | CTGCACCGCT | TCCTACATAT | GCTGCACTTG | CCACAACAGC | AATTGCTAAT | GGCGATAGGT | 3780 |
| | CAAATTTCAT | GGCAACCAAT | GCACCAATCA | AAGCAGGTAC | TGTAAATTGA | ATTGCAACGA | 3840 |
| 35 | CAACGCCTAA | TAACGTTTTA | AAAATCGGAT | GATAATCCAT | AAAGTATTTA | AAAATTTCTC | 3900 |
| - | CAAGTATCGC | ATTAGGAACT | AAACCCGCAA | CAATACCTAT | GGCGACACCT | GATAAAACTC | 3960 |
| | AAATATAAAT | ATCTTTGGGT | GTAATTGTTT | TAATTGATGT | CATAATATCA | TCCTTCCATT | 4020 |
| 10 | TATGTATATA | CATCTGTATG | CAAATAATAA | AGAGCCTTAA | GTTATAAGCT | GCCACTAGCT | 4080 |
| | TAAATTCTAA | GATGTGCATG | CCGATGTTGT | TATATTTAGG | CTAGCAGTAT | CATCTATAAC | 4140 |
| | TCAAGACTAT | GAAAAATAGT | ATATCACAAA | ATTCTGAATT | TTTAGATAAA | TAAATTGGCA | 4200 |
| 15 | ATTTTTCAAA | CATATTGTTA | CAATACACTT | TTATTTTATC | TTCATTTTTA | AAATCCATTA | 4260 |
| | ATACAATAGA | AGAAAGACAT | TCAAATGCTT | ACCAAAAAGG | TACATTATTT | GTTAGGAGCG | 4320 |
| | TATCAGCaCT | TACATATCAT | CAACACAATT | GACAATATAA | TAGAAGATAC | TGATAATAAG | 4380 |
| 50 | TGTTAAAACA | ACAGATGTTA | GGTAGTGAAC | AAATGATGGA | AAGTAAATCC | ATAGATCCAA | 4440 |
| | GAATCGTTAG | AACCAAACAA | TTGCTTGTCG | ATGCTTTTCT | TAAAATTTCT | AGAGAAAAGA | 4500 |

| | TTTACGCTCA | TTTCGCTGAT | AAAGAAGACO | CTCCTAGACTA | CACATTATCT | GTAACCATTT | 4620 |
|-------------|--------------|------------|------------|-------------|------------|------------|------|
| 5 | TAAAAGACTT | GAATGATAAT | TTGAGCATTI | CTAATGTCAT | TAATGAAAAG | GTTCTGCGTA | 4680 |
| J | ATATTTTCAT | TTCAATTGCG | AGTTATATCA | AAGATGCTGC | AAAGTCTTGC | GAATTAAATA | 4740 |
| | GTGAAGCATT | TTGCAACAAA | GCACATCAAC | GTATTAATAA | TGAATTAGAA | GATATTTTTG | 4800 |
| 10 | CGATTATGTT | AGAAAACAGC | TATCCGGAGC | ATCAACGAGA | TATCATTGTA | AATAGTGCGA | 4860 |
| | GTTTTTTAGC | AGCTGGTATC | TCAGGCTTAG | CATTACATTG | GTTTAACACG | AGTCAAGAGA | 4920 |
| | CAGCCGATGT | GTTTATCGAT | CGCAACCTTC | CATTTTTAAT | TCATCATATA | GCACATTTTT | 4980 |
| 15 | AATAAAACTT | GGTATTTAGT | CATGCATCTT | GAAATCACTA | TGTGACTTAG | GTTCATACTT | 5040 |
| | GTACACACAA | TAAAATTTAA | CGTATTACGA | TTGATTAGCC | GTGTCTAGGA | CATAAATCAA | 5100 |
| | CGTCCTATAC | TCTACAATGT | CATATTAGCA | GTCGTTAACT | GAATGAAAAT | AAGCTTGTCA | 5160 |
| 20 | TTAAAACATA | TAGATTTTAG | TGACAAGCAT | TTTTGTTTTT | GCGTACTTAA | ACAACACTTC | 5220 |
| | AGGCAATATG | TTGTTTAGGC | AACAAATGAT | ATGTGCGTGT | TTATTGGCAA | ACGTACGACA | 5280 |
| | TAGTAGTATA | GTATGTCTAA | ACAACATATG | TTGCATAGTT | GATATGCGTT | GTTTAAATAC | 5340 |
| 25 | TAAGATAGGA | GGGATTGACG | TGAGCGAGAC | AGATGAACCT | CAGGGGTTTG | AACGCACGCA | 5400 |
| | TAATATATTA | AATATTAATC | AGAGTAGTCT | GGGTGTAGTG | ACATACATTA | CAAATAAATT | 5460 |
| 30 | AAAGTCGACG | TTGAAGCAAC | ACATAATAAT | TGCTCGTGGT | AAAAAGCGAA | TCGACTATCG | 5520 |
| 30 | ACTGTCGTAT | AACTTTTACA | TACGTATTAT | GATAATGTAG | AAATCAAGAA | AATCGACTGT | 5580 |
| | GAATATACCT | ATGCTATGCC | CATTGCAATT | TTAATAAGAC | ACACGATGTC | ATTCGACAAT | 5640 |
| 35 _ | GCTCATTTCT | TTGCTCAGTT | ACGTCATCCT | GTCTTATAAA | ACAACATTGC | AGACATGTAT | 5700 |
| | ATCAAACGAC | ACTTCAATAA | CATCACTTTG | CCCATCGTAC | TACTAGTAAA | ATCGTGTCTC | 5760 |
| | AAATECCTTA | TTTTAATTCC | AAAAAtCTGC | TGGTCAAAAG | ACCGAGAAAC | TAAAAACATT | 5820 |
| 40 | ACTTAATGTG | TTGATAAATT | ACCATATAAA | AATAATCTCA | AAATATATCA | ACACTTGATT | 5880 |
| | CTAAGGAGGA | TATGACAATA | TGAAAATTTT | AGATAGAATT | AATGAACTTG | CAAATAAAGA | 5940 |
| | AAAAGTACAA (| CCACTTACTG | TAGCTGAAAA | ACAAGAACAA | CATGCATTGC | GTCAAGAcTA | 6000 |
| 45 | CTTAAGCATG I | ATCCGAGGAC | AAGTATTAAC | AACATTTTCC | ACAATAAAAG | TGGTTGATCC | 6060 |
| | AATCGGTcAG (| GATGTCACAC | CAGATAAAGT | TTATGATCTT | CGCCAACAAT | ACGGTTATAT | 6120 |
| | TCAAAATTAA t | ATTTGCTCA | CGAGGTATTG | CACTTAAGGT | GCCAACTGAC | CTCATAAACA | 6180 |
| 50 | AAGCCCATAC 1 | GATTGAAGA | CACTAATGTG | tCsaCCATGG | TGCACATTAC | GCTTCATCTC | 6240 |
| | TGTATGGGCT 1 | TATTTATTT | TCTTTTGAGA | ATTTCATTTT | AGCAGACCAA | AAAATTAAAA | 6300 |

| | TGAACGACTG | TGCCACCCGC | TTCTTTCACT | TTATTCACCA | ACTGGTCAAC | TTCTTCATTT | 6420 |
|----|-------------|-------------|-------------|------------|------------|------------|------|
| | GTGTTCACAC | CTAGAGAAAT | CATCACTTCA | TTTGGTTCAG | TATTAAGGCT | TTGCTGACTT | 6480 |
| 5 | ACATTTTGAA | AATGCTTGTn | ТТСТАТТААА | ATTACGGRTG | tTTGACCTAT | tTGAATGCCG | 6540 |
| | ACCATTTTAT | CTAACATTTG | TGGGTTTCTA | TTTATTTTAA | ATCCTAACGC | TTTATAAAAC | 6600 |
| 10 | TGTGCGCTCT | TTTCTAAATC | TTGCACATGC | AAATTAAACC | ACATTGATTG | AATCATGATT | 6660 |
| 70 | GCACCCCATT | CATTACTTAT | TATAGTTTTG | GACTTTAAGC | CAATCACTTA | ATGATAATCT | 6720 |
| | TGTTGGATTT | ATTTCAGCCA | TTAATTCAAA | GTCTACTTCA | TAACCTTTTT | CTTCCAACCA | 6780 |
| 15 | TTGCTTTTCT | GCAACACCAC | TAACAAATTC | TCCTTCTATA | ACAGTAGATT | TACCTGTCAC | 6840 |
| | TTCACTAAAA | ATTGTTGCTG | CTTCACTTAA | TGTAACTTCA | TCGGAACCAA | TCTCTATTGA | 6900 |
| | TTGATGCGTA | AAGCTTTGTG | GATGTGCAAA | AATATACGAT | GCAATTTTAG | CTATATCAAT | 6960 |
| 20 | AGAAGAAATC | ATTGTGAATT | TTATATTCGG | ATTAATAAAT | TCTGGTAATG | TAATACGTTC | 7020 |
| | ATCTTCGACT | TTAGCAATGC | GTAAAAAATT | ATCCATAAAG | AATGATGGTT | TGATAACTGT | 7080 |
| | TGCATTTATA | TTAGATTCCA | TTAATCTATT | TTCTATTTTT | GCTAGTACTT | CAAAGTGTGG | 7140 |
| 25 | GCCAGTTCGA | TTTCGATTAA | CCCCTCCCGC | AGTACTATAC | ACAATATGTT | GAATATTTTC | 7200 |
| | TTGCTCAGCT | ATTTCAATTA | TCTTCATACC | TTGTCTTAAT | TCTTCGCTAA | CATCATCTTT | 7260 |
| | AACGATTGGC | TGAATACTGT | ATAAGCCATA | CTTACCTTTC | ATCGCTGATT | GCAAACTAAC | 7320 |
| 30 | ATTATCACTC | AGATCACCTT | CArcGATTGA | TAAATGCGGA | TGTCCTATGT | CTGAAAGTTT | 7380 |
| | ACGATTATAC | TTATTTCTAG | TTAATGCACT | TACATACCAT | CCATCCTCTA | ACAACTGTTT | 7440 |
| 35 | TACAACTGCA | TTACCTTGCT | TCCCTGTTGC | GCCTATTACn | AAAATATCTT | TCAT | 7494 |
| 33 | (2) INFORMA | TION FOR SE | O ID NO: 70 |) - | | | |

(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 11802 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double

- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

| AATTTATTTC | GCCGTCCCAC | CCCAACTTGC | ATTGTCTGTA | GAAATTGGGA | ATCCAATTTC | 6 |
|------------|----------------------|------------|------------|--|--------------|-----|
| TCTTTGTTGG | GGCCCcGCCC | CAACTCGCAT | TGCCTGTAGA | ATTTCTTTTC | GAAATTCTCT | 120 |
| GTGTTGGGGC | CCCTGACTAG | AATTGAAAAA | AGCTTATTAC | AAGCGCATTT | TCGTTCAGTC | 180 |
| AATTACTGCC | ארס א מדא א מריע א מ | ССТАСАТСАТ | AGAACATTCA | THE PROPERTY OF THE PROPERTY O | ACCOTA TOTOT | 24/ |

55

40

45

| | AGCAAAGGTA ATAATGATAT TAATAATGTA CAAAAAATAT AAATCAAATC | 360 |
|---------|---|------|
| _ | ТААААСАТСА GAACCACTAA AAACAAAAAA GCACAAAATA AAATTAAATT | 420 |
| 5 | GACCACTTTT CAAAAAATC TCtTTTCaTa TTTCCACCCC TAATTTTAAT AAGCATTATT | 480 |
| | TTATATTCTC TTTTAAGTTT ATTATTCAAA AGGAAAACAG AAATATCTTT CAATATTATT | 540 |
| 10 | ATAAACATTT CAACTACTTT TAAAAACCAA CAAAAAAATA CTTATTTTAA GTAGATGAGC | 600 |
| | ATAAGTGAAC ATAGTTCTTT AGTTATAATA ATTAATTCAA CCAAAAGTCG ATTTGTTTTT | 660 |
| | GCAATTGGTT TTCATTTCCT CTTAAAGATA TTTTCATTAA ATCTGTCAAA TCAATAGACG | 720 |
| 15 | CTATATTTTT CAACTTATCT CTATATTTAT TTTTAGTACG TCTTTCTAAA TTTCCCCATT | 780 |
| | CCTCTTCTTC GTGAGTTAAT AAATGAAGCA TTGCTCGTTC TTGTATATTT TCAATCATTT | 840 |
| | TTAAATTCGG TTTTAAAATA TGCAAATCAT CAAAACAATC TTTCCAACAA TCAACCATAT | 900 |
| 20 | CTCGTTTTAA TTCAATTTCC ACACGCCATA GAAATGTTGA ATCAATTTCA ACATCTGCAT | 960 |
| | TATCTTTACG TICTTGTTTT TATTATAAAT CCGAATAAAC CTATCACTAT TACGCACACC | 1020 |
| 25 | AAAATATTTT GTTTCTGGTT TTACATTACG TCCATAAAAT ATAGTTTTCT TTACCGACTT | 1080 |
| | ATCTGACAAT GCATAATAGT CATTTAAATC AAATTCAAAA TCAAAAGCCA AATCTAATCT | 1140 |
| | CGTAAAACTA ACATCGTCCA AATAACTGAT GATATTTTGT TTTAACCAAA GCACTTCATC | 1200 |
| 30 | ATGCGAAAGC TTATTAGGAT TAAATTCAAC GCGCATALAC GTCTATTCCA AAGAGTTGCT | 1260 |
| | TTTATTTTGT CATATTCAAT ATAAACTTTT TCTTTAAGAG CTTTAGCTTT AAAGTTTGTT | 1320 |
| | TGTAAAATAT CCCAAAGCCG AATTTCAGGA TTAGTACTCA TAAAATGTGA AAGTCTCTCT | 1380 |
| 35 | GCGTTAGACA TGCTAAGATT CCCAACAATC GTTATAGCGT CAAAAGACAA TTTTGGAATA | 1440 |
| | GCTAGTGACA TCCTATGTCG ATTTAACCGG CTATTACCGG ATATTAGAGT ATCCAGTTTT | 1500 |
| 40 | ACAAATGGAT GAAACGAAAT TCAAAACACT AAAAAATATG TTCCACTAAC AGCAAAAAAA | 1560 |
| | TACCATTATG TTCCTACTAA AAAACYAAAA ATACTGGAGA ACAAATGTCA GGATATAACT | 1620 |
| | TAGGATACTA TGTAATAAAA ATTTACAATA AAAAAACAGG AAAACAAATT TCAAGTAAAA | 1680 |
| 45 | GMATACCCAT ACAAAGAGGA TAAAATAAAA AACCTCGAAC TGAAATGATG ATCTTTTCAG | 1740 |
| | CTCGAGGTTT AAATATTGGT GCCTTATTTA TATAGATTCG TTATATTATA | 1800 |
| | TCATTAACmT AATCCTTAAA GAGTTTTAAA TTAATACCTG CTAGATGATT CAAAAATGTT | 1860 |
| 50 | TCATCAACTT TTAAATAATT CAATAATTTT TGTGGTGTCA GTAAATNTCT ATCAAAATAC | 1920 |
| | AACTITAATA AACTATTCAT TTTGACAGGA CGTGACATTT CAATCACGTC GTCTAAAGAT | 1980 |
| | AATACTTTCT CGCTTTANAC AAANACAAAA ACTTACCCGA TTAAAATCAA GTAAGTTTTA | 2040 |
| 55 | | |

| | TATTTGATAA | AAAATCAATA | AGTAATTGTG | CGCCTTCAAC | TTGAATATCT | TTTACAACTG | 2160 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | GCGCGTCGAT | ATACATATCA | TACTGACCAC | CGCCTACTGC | ACGATAATTA | TTTACACAAA | 2220 |
| 5 | TTGTATATGT | CTGCTTTAAA | TCAACTGCGT | GACCTTGAAT | CATCATATTG | CTCACACGTT | 2280 |
| | GTCCCTTTGG | TCTTCCAACA | TGAATGGTAT | AACTTACGCC | ACCATATATA | TCATAATTAA | 2340 |
| 10 | AGTGTTGTGG | TTTGGGTTCA | AGGAAGTCTG | CGCTCACACT | AACTTCATCA | TTTTTCACGT | 2400 |
| | CAAAATATTC | TGCTGATCGT | TCAATGGCTT | CTTTAAGTTT | GGCACCACTT | ACAGCTAAAA | 2460 |
| | CTTTAAATGT | ATTTGGAAAT | GGGTAATTGT | TAATAACATC | TCGCATCGTC | ACGACTTGCT | 2520 |
| 15 | TGAAACCACT | AGCAGAATCA | AACAAAGCTG | TACAGGCAAC | ATCTGCGTCA | CTTTTTTCTA | 2580 |
| | ATAAAGCGTA | ATTCATAAAA | TTTGTAAAAG | GATGCGGTGC | CACACGTGCC | TCAAATGCAT | 2640 |
| | GATTAATCGT | CATATCATAT | GGCAATGTAG | TAATTTCGTA | ATCTAACCAG | TCCTCTAACT | 2700 |
| 20 | GCTTTCGTAA | ATGTTGGTCA | TCTTCATCAA | TAGTAAATGT | GGAATCATCT | ATAACAGGAA | 2760 |
| | GTAATTCACA | TGATTCAACG | GATAGATTTT | CATATTCATC | AGTACTCAAG | ACTACTCTGC | 2820 |
| 25 | CTACAGTTGT | ACCTCTCGTA | CCAGGTTGAA | TCACAGCCGT | TTGCTTAAAC | CTTTCAGCAA | 2880 |
| | TTTGTCGATG | TTGGTGACCC | GTAATAAAGA | TATCTATATC | TTTAGAAAAC | GCTTCTAACA | 2940 |
| | TGGCATATCC | TTCATTTTCA | CCCGTTAATA | CTTCGGTCGG | CGTACCACTT | TCTAAATCCT | 3000 |
| 3 0 | TTTCAAATCC | ACCATGGTAA | CAAACCACAA | TGATATCTGC | ATGTCGCTTC | ATTTCAGGTA | 3060 |
| | AGTATTGTTG | AAGTATTTCA | AAAGCACTAT | GAAACGTArT | GnCnTGAATA | TGCTCTGGTT | 3120 |
| | GTTCCCAATG | GGGAATAAAT | TGTGTCGTTA | AACCTATCAC | ACCAACAGTT | TGATCTCCAA | 3180 |
| 35 | CCTGAAAATA | CTTCACACCG | TTATCAGTCA | ATGTACTATC | ATTTTCATAT | ATATTAGCGC | 3240 |
| | ACAAAACTGG | ATAATTGAGT | CTGCGTAAAG | TGTCTTTTAA | GTATGGTAAT | CCATAATTAA | 3300 |
| 40 | ATTCATGATT | ACCAAGCGTA | CCAAAGTCGA | ATGCCATTCG | ATTATAAAA | TCAACTAAAG | 3360 |
| 40 | GCTGGCTACT | GCCGCTATGC | GCGATTAAGT | AATTACAAAA | TGGTGACCCT | TGCAAAAAAT | 3420 |
| | CACCATTATC | TATTTTAAAA | CTTTGGTCAT | ACTGCCTTCT | GTSTTGTTCT | ATAACATGAT | 3480 |
| 45 | TCGCTAGTAA | CAATCCCATA | GGTTGATATT | GATTTCTACT | CGTAAAATCT | GTTGGGAAAA | 3540 |
| | TATAACCATG | TACGTCACTC | ACGACATAAA | ATGCTATGTT | TGACATCCTC | ACTCACTCCT | 3600 |
| | TCAATCACAA | ACATCTTTCT | TATTTCTATT | ATATATTTAT | TTGAAGTCTG | TTGTAATCAA | 3660 |
| 50 | GGTTTTGTCA | CCGAGTTTTA | AACGAATCTT | TGAACCTTCC | ATACTTTCAA | GTACTTTAGC | 3720 |
| | ATTGACCTTA | ATTGTGACAT | TTCCGTTTTC | ATCTGCTTTA | ACTGTTGGCA | AAGTACTGTA | 3780 |
| | ACCTGGTGGG | TTATAATCGT | TATCTTTACT | TGAAAATTGT | CCGATTTGAC | GTCCGCCTTC | 3840 |

. 55

| | TATIGICATI TCAAATGGCT CATTTACAGA AACATTTTGC GGGATATCAA ATGTTACTTT | 3960 |
|----|---|------|
| 5 | TTCGTTCTGA TTTGGTGGTG TATGATCATC TGGTGTGTTT GGCTGAGGAT CTGCGCCTTT | 4020 |
| | TTCGCTGCCA TAACTACCTG CTTTAAATGT TGTTGGATCA TACCATTTAT AACCACTCGG | 4080 |
| | CGGTTGTGAC CATGGCTCTT TTTCAGGCTC AGTTGAACGC TCTGGTCGTT CAAAATCAAG | 4140 |
| 10 | CAACTTAGTC TITGTATCTA ATGTTAGGCT ACTCGCCTTA AGTGATTTCC CATCATTATC | 4200 |
| | TTTAGACATC CAAGCCGTTA TATTATTTAA TAGCTTACCG TTGTCTTGTT CTTTAAAACC | 4260 |
| | ATCATATGTT TTCTTCTTTT CTCCATTATC TTCTCTTACA TATTTGGGCG AACTATCTTC | 4320 |
| 15 | CACAAGTGAT GAATCACCGA TAAATGCTGC TTTACCTTTT CCAACTTTAG AAATTGCTAC | 4380 |
| | ATAGGGGCCT TCTGCTTTAC CGCCCCCATT ATAAATACCT TGATCTACAG CATGTGACCA | 4440 |
| | TITACTITIC GCTGGCAATT GTTCTGGTGT ATACACAATA CCTTTTGCTT TCTCTGGATT | 4500 |
| 20 | AGTAATTGCT AATGTCGATC CGGCATGCAT AGAGACAGAT TTCACACCTT CAGTAATACC | 4560 |
| | GAAACTTTCT TTTGAAGAAA CAATATTGCT CGTATTTAAA TCACCTAGTG CATTATATCG | 4620 |
| 25 | AAAACGTACG CCAAAGTTTG TAGATAACCA ATCTGAACTT TTCACACCTT GCATTGCAGT | 4680 |
| | AGAACTTTTT TCTTCTGCAT TCATACCTTT CGACATATCT TCATATGCTC CACGTCGATA | 4740 |
| | ACCATTCATT GCCTCCGATG AATCAATACG ATTTAAATTT CGGTCAGCAT TGTAATGATC | 4800 |
| 30 | TGAAATAAAG ACAACATTGC CACCTTGTTL CACATATTTA ACAATTGCTG CCTGTTCTGA | 4860 |
| | TTCTTTGAAA GGAATGTTAG CCTCAGGAAT TACAAATATT TTGGAACTTT TCAAACTTGC | 4920 |
| | TTCTGTTATG TTCGAATGAC CATCAATAGC TTTAACGTCA TAACCTTGTT TTTGTATTGA | 4980 |
| 35 | ATCCGCATAA TCTGAAAATG CACCATCACT AACCCAATCT GCAGCACCAG CTGTTTGACC | 5040 |
| | ATGAGAACGA TCGAATAATA CCGTTCGCTG TTGCTTTGTA GGTTGCGATT CATGCGTTAT | 5100 |
| 40 | AGCTAAAGAT TGCGGTAAAG CACTTAATGA TACCGTTGCA ACAATTGCAG AGACAGTTAA | 5160 |
| | TGACTTATAT ATTTTTTCA TTTTGTGAGG CTCCTTTTAA AATAAATTTG TTCTTGAATT | 5220 |
| | ATAGGATAAA AATTCGTTGC ATATGAGCAA TTTAACGAAA AATTTACAAA ATCTTATCAA | 5280 |
| 45 | ACTCTTAAAG AAAGTTATTA AAATTCATTT TTATAAAATA CTTTTTAACA TTTAAATGTG | 5340 |
| | GTACGCTATA AGTGTAATTT CATTGCATAC ATATTACACG ATTAAGAATG TGAAGGGGAC | 5400 |
| | AGTTATCAAA TGAAAAATTT TAAGTGTTTA TTTGTATTAA TGTTAGCAGT CATTGTTTTT | 5460 |
| 50 | GCAGCAGCAT GTGGAAACTC AAGTTCTTTA GATAATCAAA AGAACGCTAG TAATGATTCG | 5520 |
| | GATTCTAAAT CAGGAGGATA CAAACCTAAA GAATTAACCG TTCAATTTGT ACCTTCGCAA | 5580 |
| 55 | AATGCTGGAA CATTAGAAGC TAAAGCAAAA CCATTAGAAA AATTACTATC TAAAGAATTA | 5640 |
| , | | |

| | TCTAAAAAAG | TIGATGTIGG | TTTCTTACCA | CCAACGCCAT | ACACATTAGC | ACATGATCAA | 5760 |
|----|------------|------------|------------|------------|------------|------------|---------------|
| | AAAGCAGCTG | ATTTATTATT | ACAAGCACAA | CGTTTCGGTG | TAAAAGAAGA | TGGTTCAGCA | 5820 |
| 5 | AGTAAAGAAC | TTGTAGATAG | TTATAAATCA | GAAATTCTTG | TTAAAAAAGA | CTCAAAAATT | 5880 |
| | AAAAGCTTGA | AAGATTTAAA | AGGTAAGAAA | ATTGCCTTAC | AAGATGTAAC | ATCAACTGCT | 5940 |
| 10 | GGATATACAT | TCCCACTTGC | GATGTTAAAA | AACGAAGCAG | GTATTAATGC | AACTAAAGAT | 6000 |
| | ATGAAAATTG | TGAATGTTAA | AGGTCATGAC | CAAGCAGTTA | TCTCATTATT | AAATGGAGAt | 6060 |
| | GTAGATGCTG | CGGCTGTATT | TAACGATGCA | CGTAATACTG | TGAAAAAAGA | CCAACCAAAT | 6120 |
| 15 | GTATTTAAAG | ACACACGAAT | TTTAAAATTA | ACACAAGCTA | TTCCGAATGA | CACAATTTCT | 6180 |
| | GTAAGACCAG | ATATGGATAA | AGATTTTCAA | GAAAAATTGA | AAAAAGCTTT | TATAGACATT | 6240 |
| | GCTAAATCAA | AAGAAGGTCA | CAAAATTATT | AGCGAAGTTT | ATTCACATGA | AGGATACACA | 6300 |
| 20 | GAAAEGAAAG | ATTGAAATTT | CGACATTGTA | AGAGAGTACG | AAAAATTAGT | TAAAGATATG | 6360 |
| | AAATAATCAT | TATTTAACAA | ATGAATCATT | AGCGAATTTG | GTATTAAAAG | CTTTCGTTCA | 6420 |
| 25 | ATAGATATAT | TCTAGATTAA | TATTGAAAAG | CTAGGCGCTA | AACTGAAACA | GATATAGAAA | 6480 |
| 23 | GGTGTCGCTG | TACATTTGAA | ACCATTTGTA | CACAGAAACC | CAATGTCTAT | GATATTTCAG | 6540 |
| | TTTACCTTGG | CTTTTCTTTA | TTAAAGAAAG | GTGTCAAACA | TGAGTCAAAT | CGAATTTAAA | 6600 |
| 30 | AACGTCAGTA | AAGTCTATCC | TAACGGTCAT | GTAGGCTTGA | AAAATATTAA | CTTAAATATT | 6660 |
| | GAAAAAGGTG | AATTTGCAGT | TATTGTCGGA | CTATCTGGTG | CTGGGAAATC | CACGTTATTA | 6720 |
| | AGATCTGTAA | ATCGTTTGCA | TGATATCACG | TCAGGTGAAA | TTTTCATCCA | AGGTAAATCA | 678C |
| 35 | ATCACTAAAG | CCCATGGTAA | AGCATTATTA | GAAATGCGCC | GAAATATAGG | TATGATTTTC | 68 à 0 |
| | CAACATTTTA | ATTTAGTTAA | ACGGTCAAGT | GTATTACGAA | ATGTACTAAG | TGGACGTGTA | 6900 |
| | GGTTÄTCACC | CTACTTGGAA | AATGGTATTA | GGTTTATTCC | CAAAAGAAGA | CAAAATTAAG | 6960 |
| 40 | GCAATGGATG | CACTAGAACG | CGTCAATATC | TTAGATAAAT | ATAATCAACG | CTCTGATGAA | 7020 |
| | TTATCAGGTG | GCCAACAACA | ACGTATATCT | ATTGCACGTG | CGCTATGCCA | AGAATCTGAA | 7080 |
| 45 | ATTATTCTTG | CAGATGAACC | AGTTGCTTCA | TTAGACCCAT | TAACTACGAA | ACAGGTTATG | 7140 |
| | GATGATTTAA | GAAAAATCAA | CCAAGAATTA | GGCATCACAA | TTTTAATTAA | TTTACATTTT | 7200 |
| | GTTGACTTGG | CAAAAGAATA | TGGCACACGC | ATCATTGGTT | TACGTGATGG | TGAAGTTGTC | 7260 |
| 50 | TATGATGGTC | CTGCATCTGA | AGCAACAGAT | GACGTATTTA | GTGAAATATA | TGGACGTACA | 7320 |
| | ATTAAAGAAG | ATGAAAAGCT | AGGAGTGAAC | TAACATGCCT | TTAGAAATAC | CTACAAAGTA | 7380 |
| | TGACTCCCTT | TTAAAGAAAA | AGGTTTCTTT | AAAAACGAGT | TTTACCTTCA | TGTTAATCAT | 7440 |

| | AATACCTCAA ATAGGTGATC TATTCAAACA AATGATTCCA CCTGATTTCG AGTATTTACA | 7560 |
|-----------|--|------|
| - | ACAAATTACA ACGCCAATGT TAGATACCAT TCGAATGGCT ATCGTAAGTA CAGTATTAGG | 7620 |
| 5 | TAGCATCGTT TCAATACCAA TTGCGTTATT ATGTGCTAGC AATATCGTTC ATCAAAAGTG | 7680 |
| | GATTTCAATA CCCTCGCGCT TTATTTTAAA TATAGTTCGT ACTATTCCAG ATTTGTTATT | 7740 |
| 10 | AGCAGCAATC TTTGTGGCTG TATTTGGAAT CGGTCAAATT CCAGGGATAT TAGCACTGTT | 7800 |
| | TATTTTAACT ATCTGTATTA TTGGAAAATT ATTATATGAA TCATTGGAAA CGATAGATCC | 7860 |
| | AGGTCCAATG GAAGCAATGA CGGCTGTTGG CGCTAATAAA ATAAAATGGA TTGTTTTCGG | 7920 |
| 15 | TGTTGTACCA CAAGCCATAT CGTCATTTAT GTCATACGTA TTATATGCAT TTGAAGTAAA | 7980 |
| | TATACGTGCT TCAGCTGTGC TTGGATTAGT CGGCGCTGGC GGTATTGGAT TGTTTTATGA | 8040 |
| | TCAAACACTT GGTTTATTTC AATATCCAAA AACAGCAACG ATTATTTTAT TTACTTTAGT | 8100 |
| 20 | TATCGTCGTC GTCATTGATT ACATCAGTAC GAAAGTGAGG GCACATCTCG CATGACACAG | 8160 |
| | GAAATAGCAA AATATAATGT TCACACAAAA GCACACAAAC GAAAATTGAT TAAAAGATGG | 8220 |
| | CTTATTGCAA TTGTCGTCTT AGCTATTATC ATCTGGGCAT TTGCAGGTGT ACCAAGTTTA | 8280 |
| 25 | GAACTTAAAA GTAAATCATT AGAAATCTTA AAATCCATAT TCAGCGGATT ATTCCATCCT | 8340 |
| | GATATCAGCT ATATCTATAT ACCAGATGGC GAAGACTTAT TACGTGGTTT ACTTGAAACC | 8400 |
| 30 | TTTGCGATAG CCGTTGTAGG TACTTTCATC GCCGCAATTA TCTGTATTCC ATTAGCATTT | 8460 |
| | CTAGGTGCAA ATAATATGGT AAAGCTACGC CCAGTTTCAG GTGTTAGCAA ATTTATTTTA | 8520 |
| | AGTGTTATAC GTGTCTTCCC AGAAATTGTA ATGGCACTTA TATTTATCAA AGCTGTTGGC | 8580 |
| 35 | CCAGGTTCAT TTTCAGGTGT ATTAGCTTTA GGTATCCATT CCGTAGLATG CTTGGGAAAC | 8640 |
| | TTTTAGCTGA AGATATTGAA GGTCTAGATT TCAGTGCTGT AGAATCATTA AAGGCCAGTG | 8700 |
| | GTGCGAATAA GATTAAAACA CTCGTATTTG CAGTCATACC ACAAATTATG CCTGCCTTTC | 8760 |
| 40 | TATCACTCAT ACTITATEGE TITGAACTAA ACTITACGITE AGCITETATA CIGGGGETAA | 8820 |
| | TTGGGGCTGG TGGTATCGGG ACACCACTCA TATTTGCCAT TCAAACACGT TCTTGGGACC | 8880 |
| | GTGTAGGTAT TATATTAATC GGTTTAGTAC TAATGGTCGC AATTGTCGAT TTAATTTCCG | 8940 |
| 45 | GTTCAATCCG AAAACGTATT GTTTAACATT AAATCAGGAT ACTCCTAAAT AAGAAGTCCT | 9000 |
| | ACCGTCTTAC GTTTCTCTAT TATAATAAAA ACAGCAGTGA AGAAAACTAT TGTTATAGTT | |
| 50 | AACTTCACTG CTGTTTTTAT AATATCTAAA TTTATTCTAT TTCAATTCCT TTAAATAACT | 9060 |
| <i>50</i> | TTTACCGAAC TCTGGTAATG TTACGTTGAA ATTATCTGCT ATAGTTGCAC CGATAGAACT | 9120 |
| | GAATGTAGTA TCACTTTCTA GTGCATGACC ACCTTTAAAT TTCGGACTGT ACATAATTAC | 9180 |
| | · ···································· | |

| | TGTAATAATT | ACTAAATCGT | CTTCTTTTAA | GTTGCTAAAC | AGTTCTGGCA | AGCGATCATC | 9360 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | GAAATCTTTA | ATTGCTTGTG | CATAACCTGG | TTTATCACGA | CGATGACCGT | ATAATGCATC | 9420 |
| 5 | AAAGTCTACT | AAGTTTAAGA | AGCTAATACC | TGTGaAATCT | TTCTTAACAA | TTTTCATCAA | 9480 |
| | TTGATCCATA | CCGTCCATGT | TACTCTTCGT | ACGAACCGCT | TCTGTTACAC | CTTCACCATC | 9540 |
| 10 | ATAAATGTCA | TTAATTTTAC | CGATGGCAAT | AACATCATAA | CCACCGTCTT | TCAAATGATC | 9600 |
| | TAAGACAGTT | TTACCAAAAG | GTTTTAACGC | ATAGTCATGT | CGATTAGATG | TACGTGTAAA | 9660 |
| | GTTTCCTGGT | TCACCAACAT | ATGGACGTGC | GATAATACGA | CCAATTAAAT | ATTTAGGGTC | 9720 |
| 15 | TTTTGTCAAC | TCACGAACCT | TTTCACAAAT | ATCATATAAC | TCTTCTAATG | GGATAATGTC | 9780 |
| | TTCATGTGCA | GCAATTTGCA | ATACTGGGTC | TGCACTTGTA | TAAACAATTA | AGTCACCAGT | 9840 |
| | TTTCATTTGG | TGCTCGCCCC | ACTCATCGAT | AATTTGCGTA | CCCGATGCCG | GTTTGTTAGC | 9900 |
| 20 | AACAACTTTA | CGACCTGTCA | TTTCTTCAAT | TTGTTGAATT | AACTCTTCAG | GGAATCCATT | 9960 |
| | AGGGTATACT | TTAAAAGGTT | GCATAATATT | TAATCCCATA | ATTTCCCAGT | GACCAGTCAT | 10020 |
| | TGTATCTTTA | CCAACTGAAG | CTTCACTCAA | TTTAGTATAG | TATGCTTCTG | GTTGTTCAAC | 10080 |
| ?5 | TGCATTTACT | ACTGGTAATT | TATCGATGTT | CCCTAGACCT | AACTTTTCAA | GGTTTGGTAA | 10140 |
| | AGTTTGATCG | AAACCTTCTA | AGGTATGTCT | TAAAGTATGT | GAACCTTCAT | CTTTAAAATC | 10200 |
| 30 | AGCTGCGTCT | GGCGCTTCAC | CAATACCTAC | TGAATCCATT | ACGATTAAAT | GTACACGATT | 10260 |
| | AAATGGTCTT | GTCATAGCTA | TCACTCCCAA | AATTTATATA | TATTAGTAAT | CTGAATCTGC | 10320 |
| | TTCTAAACCT | TGCATAATTT | GAACACCTGC | GCTCGCACCA | ATACGTGTCG | CACCTGCTTC | 10380 |
| 35 | AACCATTTTA | TTGAAATCTT | CTAAATTACG | TACGCCACCT | GATGCTTTTA | CTTCTACATC | 10440 |
| | AGCACCTACT | GTATCTTTCA | TTAATTTAAC | GTCTTCTGCA | GTCGCACCGC | CACCTGCAAA | 10500 |
| | ACCTGTTGAA | GTTTTAACGA | AGTCCGCACC | AGCCGCTTTT | GTTAATTCAC | TCGCTTTTAC | 10560 |
| 10 | AATTTCGTCA | TGGTCCAACA | ATACCGTCTC | AATAATCACT | TTTACTGTGT | GACCTTTCGC | 10620 |
| | AGCTTTAACC | ACTGCTTCAA | TGTCTTGTTG | TACATCATCA | AAACGTCCAT | CTTTTAATGC | 10680 |
| | GCCGATGTTG | ATGACCATGT | CAATTTCATC | TGCACCATTT | TGAATTGCAT | CTTCTGTTTC | 10740 |
| 15 | AAATGCTTTC | GTTGCAGTTG | TCGACGCACC | TAATGGGAAT | CCTATTACCG | TACAAACGAG | 10800 |
| | CACCTCTGAA | TCAGCTAGTC | GCTCTGCTGC | ATATTTAACA | TGTGTTGGAT | TCACACATAC | 10860 |
| 50 | AGATTTAAAA | TTGTATGctT | TCGCTTCATC | GATGATTTGA | TCGATTTGCG | TACGTGTTGA | 10920 |
| | CTCAGGCTTC | AATAAAGTGT | GATCTATATA | TTTCTCAAAT | TTCATACTTA | CTACTCCTCG | 1098 |
| | TGTTATATAA | TCTCTTTATT | TAATTTTACT | ATAAATACGA | ATATATCTCG | CGAATTTATA | 11040 |

| ATACTCATTA | AACCTAAAAT | AATTAAAATA | ATACCGAAAT | GTGAACTTAA | TGCATCATTG | 11160 |
|--------------|-------------|-------------|------------|------------|------------|-------|
| CCTGGGAAAT | TTAATGCTTT | AAAATCGATT | AGAGCCGCAG | CAATCGCAAT | ACCTACAGAT | 11220 |
| ACCGCCACAT | TAATAATTAA | AAAAATATTA | CCAATAGCCA | CACCTGTCAT | ATTAAGATCT | 11280 |
| ATTGTTTTAA | TGGCTTCGTT | AAGTAAAGGT | GCATACATTA | AAGCAAAGCT | ACCTGCAAAG | 11340 |
| AATATCATAG | AAATGACGAA | GATTGAAATG | TGATTACCTA | CTGCAAATGC | AGGTAAAATC | 11400 |
| AAGCTCAGTG | CTATTAAAAT | AATTGCTGTG | ATAATCGCTT | GTTTTGAATT | CAGATATTCG | 11460 |
| CCGATTTTAC | CACTTAGTGC | ACCAACAATG | ACTGCTACTA | TATAACCCGG | TACTAATAAC | 11520 |
| AGTGATGTTG | TGTCTAGTTG | CAGATGATAA | ATTTGCTCCA | TTATGAATGG | GAACGTAAAA | 11580 |
| ATATAACCCA | ATTGGATAGC | ATACATTACA | ААТАСТАТАА | ATAAAAATGA | AGCATAACGT | 11640 |
| TTATTTTGGA . | AAAATGATTT | ATTTACTAAT | GGACGTTGCG | CATTTTTAAT | ATATAGCGCA | 11700 |
| AAAACGATAA | TCGCAATTAA | GGCACCAATC | ATATATAACC | AATTAAAGTT | CGTAATAAAC | 11760 |
| AGCATGACTG ' | TTGTAGCAGG | GGATCCTCTA | GAGTCGAnCC | TG | | 11802 |
| (2) INFORMA | TION FOR SE | Q ID NO: 71 | : | | | - |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1196 base pairs
- (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

| 60 | GACGAAATCG | AGCTTTAATT | AACAAGTTCA | GATGTTGATA | IGCGAAACAA | CIANAGANGA |
|-----|------------|------------|------------|------------|------------|------------|
| 120 | ATTAATCAAA | TAAAGATCGT | AACAAGCACT | GATAAGGAAA | AAATCTAACA | ATCAAAATCC |
| 180 | ATTGAACAAG | AAAAGAAGCA | ATGCGATGAC | GACATTAACA | AGGTCATAAC | TACTTCAACA |
| 240 | GCTAAAGAAG | TTTAGTGAAA | ACATCAAAGA | gCATTGCAAG | TTTAGCGCAA | CAAAAGAACG |
| 300 | GATCAAAATC | TGACGAAATC | AAGCTTTAAT | AAACGTGTAC | TGATATTGAT | ATGCGAAAAA |
| 360 | ATACTTCAAC | AATTAATCAA | TTAAAGATCG | AAACAAGCAC | AGATAAGGAA | CAAATCTAAC |
| 420 | GCAAAAGCAC | AATTGAGCAG | CTAAAGAAGA | AATGCGCTGA | CGACATTAAC | AAGGTCATAA |
| 480 | GATGCGAAAA | AGCTAAAGAA | ATTTAGTGAA | GACATCAAAG | AGCATTGCAA | AACTTGCACA |
| 540 | CCAGATTTAA | CAATTCAAAT | GTGATCAAAT | AATGCGAAgc | AGCCTTAGCT | ATGCAATAAA |
| 600 | CGAGCACTAC | AGCTGAAAAA | AAATTGACGA | GCGCTCAAAG | AAAAGCAAAA | CACCTGAGCA |
| 560 | TTAGGTTTAG | AGGATTAAAC | AATTAAATCG | ACTATAGATC | GAATGCTCAA | AAAACGTTGA |

| TTG | AAGCAAC | ACCTGAGCAA | ATCCTAGTTA | ATGGTGAACT | CATTGTACAT | CGTGATGACA | 780 |
|------|---------|--------------|--------------|------------|------------|------------|------|
| TCA | TTACAGA | ACAAGATATT | CTTGCACACA | TAAACTTAAT | TGATCAGCTT | TCAGCAGAAG | 840 |
| TCA | TCGATAC | ACCATCAACT | GCAACGATTT | CTGATAGCTT | AACAGCAAAA | GTTGAAGTTA | 900 |
| CAT | TGCTTGA | TGGATCAAAA | GTGATTGTTA | ATGTTCCTGT | AAAAGTTGTA | GAAAAAGAAT | 960 |
| TGT | CAGTAGT | CAAACAACAG | GCAATTGAaT | CAATCGAAAA | TGCGGCACAA | CAAAAGATTA | 1020 |
| ATG. | AAATCAA | TAATAGTGTG | ACATTAACAC | TGGAACAAAA | AGAAGCTGCA | ATTGCGnAAG | 1080 |
| TTA | ATAAGCT | TAAACAACAA | GCAATTGGAT | CATGTTNAAC | AATGGCACCT | GGATGTTCCA | 1140 |
| TTC | agttgaa | GGAAATTTCA | ACAACAAGGA | ACAAGCGCCn | GATTGGAACA | ATTTGA | 1196 |
| (2) | INFORM | ATION FOR SE | EO ID NO: 72 | 2 : | | | |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1519 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

| CAA | TCGTTTC | AACGCTATTA | TCTTTAGACA | ACAATTGTAA | GCGTGTATGT | GCAGTTTCTA | 60 |
|-----|---------|------------|------------|------------|------------|------------|-----|
| AAC | AGTCTAT | AATTCGAGTT | CTTAATTCAG | CTGGATCATC | TTTAAAAATA | AAATCCATCG | 120 |
| CTG | CAACTTT | GTAGACAAAT | GTTAAATAGG | TAAGTTCACT | GTGACTCGTA | ACGAAAATAA | 180 |
| TGT | TACCAAC | TGGGTCATGC | TTACGAATTT | CACTGCCTAA | TTTGATACCA | TTAATATCAG | 240 |
| TTG | AAAGTTG | AATATCTAAA | AAGTAACAGC | CTATGTCATT | CATATTTTTA | GCTTGCTCAA | 300 |
| GCA | CCTCATA | AGGATTATCA | GTTGCGAGGG | CAATTTCCAT | AGGCTTTTCT | TCTATCATTA | 360 |
| TAT | ÄATTTTŢ | AATAATGGTA | ACCATGTTTT | CTCTTTGTTT | TGGATCGTCT | TCGCAAATGA | 420 |
| AAA | TTTTCAT | ACATTCACAT | CCTTATGGCT | AGTTGTTAAT | AATTTCAACT | TTTTGAATAA | 480 |
| AGA | AACCATT | TTCGATAATT | GTATCTAATA | AGACATTGTC | TGCATTATCA | GCAATTTCTT | 540 |
| TTA | AAGTTGA | TAGACCTAAA | CCACGACCTT | CACCTTTAGT | AGAAAAACTT | TCTTGGAACA | 600 |
| ATI | CATGAAT | GCGTGGTATA | TCATCAGCGC | ATTTATTCAT | AACAATAAAC | GTTACTGAAT | 660 |
| TTI | CACTITC | AATAAATGCA | ACGCGAATGA | TAGGGTCATC | AATTTCAGTT | GATGCCTCAA | 720 |
| TTG | CATTATC | AAGAATAATA | CCAATACTGC | GACTTAAATC | GATCATATTC | AAGTTAATGC | 780 |
| TAC | TTACTTC | ATCGGGTATT | TCGATACTAA | TCGGAATATT | CATTTCTTGT | GCACGTAAAA | 840 |
| TTI | TCGCAGT | AATTAAGCCT | TTAATTTCAC | GTACTTTAAG | ATTCTCGATA | CCATTTAATT | 900 |

| GTAGGCCAGG | CATGTCATCT | TCTCGAATGT | ATTCTGAAAG | TGTCGTTAAG | ATATTGACAT | 1020 |
|------------|------------|------------|------------|------------|------------|------|
| AATCATGACG | GAACTTGCGC | ATTTCGTTGT | TGATAGCTTC | AATCTTCAAT | GTATATTCAT | 1080 |
| AATAGGTTTC | AATTTCTTCT | TGATTACGTT | TATATTTCAT | CTCTTTAAGG | AGAAATTGAG | 1140 |
| AAATAACAAA | TGTTAATATA | СТТАААААТА | TAGTGATACC | AATAAAAATA | AAAGAATACT | 1200 |
| GCCTTATTAC | TTTAGCTTCA | TCCGAGTTTA | TTTGTGAATA | AAAGAAAAAT | AATGAAAAAG | 1260 |
| TAAGCAGTAA | GATAGTCGAA | ATAACTATTA | AAAATCCTTT | GTTTAGTATT | AGATATGGTG | 1320 |
| TGCTAATTTT | TTTGAGAACT | CTATTTATTA | TATATGAGAA | TAGTATACTA | ATAGTCACAT | 1380 |
| AAACTACAAA | AAAGCTAGGG | AATATTACAA | ATATACTATC | AGAAATTTTG | GTGGATATAT | 1440 |
| GCATATATAA | CTATATACCT | GTAGTTAGCA | CnGTnATAGG | AATAATCnGG | CGAGGTCCAT | 1500 |
| AATCCACCAA | AATAGAATA | | | | | 1519 |

(2) INFORMATION FOR SEQ ID NO: 73:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5445 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

| | GTAGGAATCT | CTTTGTCTTT | TTGGGAGGAC | ATTTAATATG | AATGTATATT | TAGCAGAATT | 60 |
|---|------------|------------|------------|------------|------------|------------|------|
| | CCTAGGAACT | GCAATCTTAA | TCCTTTTTGG | TGGTGGCGTT | TGTGCCAATG | TCAATTTAAA | 120 |
| | GAGAAGTGCT | GCGAATGGTG | CTGATTGGAT | TGTCATCACA | GCTGGATGGG | GATTAGCGGT | 180 |
| | TACAATGGGT | GTGTTTGCTG | TCGGTCAATT | CTCAGGTGCA | CATTTAAACC | CAGCGGTGTC | 240 |
| | TTTAGCTCTT | GCATTAGACG | GAAGTTTTGA | TTGGTCATTA | GTTCCTGGTT | ATATTGTTGC | 300 |
| | TCAAATGTTA | GGTGCAATTG | TCGGAGCAAC | AATTGTATGG | TTAATGTACT | TGCCACATTG | 360 |
| - | GAAAGCGACA | GAAGAAGCTG | GCGCGAAATT | AGGTGTTTTC | TCTACAGCAC | CGGCTATTAA | 420 |
| | GAATTACTTT | | | | | | 480 |
| | TTTATTTATC | | | | | | 540 |
| | AATTGTTGCA | | | | | | 600 |
| | | | | | | GTGGTTCAAA | 660 |
| | TTGGTCATAT | | | | * | | 720 |
| | GGTATACGCT | | | | | | 780 |
| | | | | | | | , 60 |

55

5

10

15

20

25

30

35

40

45

| | CGAATCAATT | TACTAAAATA | AAAAGAAACG | TAAATAGCAT | AATTTAACAT | GTTTGATTCA | 900 |
|----|------------|------------|------------|------------|------------|-------------|------|
| | TGGATTATGC | TATTTTTTCG | CCAAAATTTA | ACAGATTTTG | TACAATGGGT | TAGCGATTAT | 960 |
| 5 | TTTTTAATAA | AGGAGATACT | ACTAATGGAA | AAATATATTT | TATCTATAGA | CCAAGGAACA | 1020 |
| • | ACAAGCTCAA | GAGCGATTTT | ATTCAATCAA | AAAGGGGAAA | TTGCAGGGGT | AGCACAACGT | 1080 |
| _ | GAGTTTAAGC | AATATTTTCC | ACAATCAGGT | TGGGTTGAAC | ATGATGCAAA | TGAAATTTGG | 1140 |
| 0 | ACATCTGTGT | TAGCTGTAAT | GACGGAAGTA | ATTAATGAAA | ATGATGTTAG | AGCTGATCAA | 1200 |
| | ATTGCAGGTA | TCGGTATTAC | AAACCAACGT | GAAACAACGG | TTGTTTGGGA | CAAaCATACT | 1260 |
| 5 | GGCCGCCCAA | TTTATCACGC | AATTGTTTGG | CAATCACGTC | AAACACAATC | AATTTGTTCA | 1320 |
| | GAATTAAAAC | AACAAGGATA | TGAACAAACA | TTTAGAGATA | AGACAGGATT | ACTTTTAGAT | 1380 |
| | CCGTATTTTG | CAGGTACAAA | AGTTAAATGG | ATTCTAGACA | ATGTTGAAGG | TGCACGAGAA | 1440 |
| 20 | AAAGCAGAAA | ATGGCGATCT | ATTATTTGGA | ACGATTGATA | CTTGGTTAGT | ATGGAAATTA | 1500 |
| | TCaGGaAAAg | CEGCGCATAT | TACTGATTAT | TCaAATGCGA | GTCGTACATT | AATGTTTAAT | 1560 |
| | ATCCATGATT | TAGAATGGGA | CGATGAGTTA | TTAGAACTAL | TACAGTACCT | AAAAATATGT | 1620 |
| 25 | TGCCAGAAGT | TAAAGCTTCG | AGTGAAGTAT | ATGGTAAGAC | AATTGATTAC | CACTTCTATG | 1680 |
| | GTCAAGAAGT | ACCAATCGCT | GGAGTAGCTG | GTGATCAACA | AGCAGCATTA | TTTGGACAAG | 1740 |
| | CTTGCTTCGA | ACGTGGTGAC | GTGAAAAACA | CATATGGAAC | TGGTGGCTTC | ATGTTAATGA | 1800 |
| 10 | ATACAGGTGA | CAAAGCGGTT | AAATCTGAAA | GTGGTTTATT | AACAACAATT | GCTTATGGTA | 1860 |
| | TTGATGGAAA | AĞTAAATTAT | GCGCTTGAAG | GTTCCATCTT | TGTTTCGGGT | TCAGCAATCC | 1920 |
| 15 | AATGGTTACG | TGATGGATTA | AGAATGATTA | ATTCAGCACC | ACAATCAGAA | AGTTATGCGA | 1980 |
| | CACGAGTTGA | CTCTACTGAG | GGTGTTTATG | TTGTTCCAGC | TTTTGTAGGT | TTAGGAACAC | 2040 |
| | CATATTGGGA | TTCTGAAGCA | CGTGGTGCGA | TTTTCGGTTT | AACACGTGGA | ACTGAAAAAG | 2100 |
| 10 | AGCACTTTAT | CCGTGCAACT | TTAGAATCAC | TATGTTACCA | AACTCGTGAC | GTTATGGAAG. | 2160 |
| | CAATGTCAAA | AGACTCTGGT | ATTGATGTCC | AAAGTTTACG | TGTCGATGGT | GGTGCAGTTA | 2220 |
| | AAAATAACTT | TATTATGCAG | TTCCAAGCAG | ACATTGTTAA | TACTTCTGTT | GAAAGACCTG | 2280 |
| 5 | AAATTCAAGA | AACTACAGCT | TTAGGTGCTG | CATTTTTGGC | AGGTTTAGCA | GTTGGATTCT | 2340 |
| · | GGGAGAGTAA | AGATGATATC | GCTAAAAACT | GGAAATTAGA | AGAAAAATTC | GATCCGAAAA | 2400 |
| | TGGATGAAGG | CGAAAGAGAA | AAATTATATA | GAGGTTGGAA | AAAAGCTGTT | GAAGCAACAC | 2460 |
| io | AAGTTTTTAA | AACAGAATAA | ACTTGTAGAT | TAGACTTTTG | TATAAACATT | GTGATACAAT | 2520 |
| | СААТТТААСТ | ТААТАТТТСА | ATCGAGAAGC | GAGAGATTTG | TTCGAACATG | TACAATTCAA | 2580 |

| | GCATTGTCTA CTTTTAAGAG AGAACATATT AAAAAGAATT TAAGAAATGA TGAATATGAT | - 2700 |
|-----------|---|-----------|
| | TTAGTAATTA TTGGTGGCGG TATTACAGGT GCAGGTATTG CACTAGACGC GAGTGAAAGA | 2760 |
| 5 | GGAATGAAAG TTGCATTAGT TGAAATGCAA GACTTTGCAC AAGGAACAAG CTCAAGATCT | 2820 |
| | ACAAAATTAG TCCATGGTGG TTTGCGTTAC TTAAAACAAT TCCAAATTGG AGTAGTTGCC | 2880 |
| | GAAACTGGTA AAGAACGTGC GATTGTTTAT GAAAATGGGC CTCATGTTAC GACTCCAGAG | 2940 |
| 10 | TGGATGCTTT TACCAATGCA TAAAGGTGGA ACATTTGGTA AATTCTCAAC ATCAATTGGT | 3000 |
| | TTAGGAATGT ATGATCGTTT AGCAGGTGTT AAGAAGTCTG AACGTAAAAA AATGTTATCT | |
| 15 | AAAAAAGAAA CTTTAGCTAA AGAACCATTA GTTAAAAAAG AAGGTCTAAA AGGCGGCGGT | |
| | TACTATGTTG AATATCGTAC TGACGATGCG CGTTTAACTA TTGAAGTTAT GAAGCGTGCT | 3120 |
| | GCTGAAAAAG GCGCAGAAAT TATCAACTAT ACTAAATCTG AACACTTCAC TTATGATAAA | 3180 |
| 20 | AATCAACAAG TAAATGGTGT TAAAGTTATA GATAAATTAA CTAATGAAAA TTATACAATT | 3240 |
| | AAGGCTAAAA AAGTGGTTAA TGCAGCAGGT CCATGGGTTG ATGATGTTAG AAGTGGTGAT | 3300 |
| | TATGCACGCA ATAATAAAAA ATTACGTTTA ACTAAAGGTG TACATGTTGT TATTGATCAA | 3360 |
| 25 | TCAAAATTCC CATTAGGTCA AGCAGTATAC TTTGATACTG AAAAAGATGG AAGAATGATT | 3420 |
| | TTTGCAATTC CACGTGAAGG AAAAGCGTAT GTAGGTACTA CAGATACATT CTATGACAAT | 3480 |
| | ATCAAATCTT CACCATTAAC TACACAAGAA GACAGAGACT ATTTAATCGA TGCGATTAAT | 3540 |
| 30 | TACATGTTCC CTAGTGTTAA TGTTACAGAT GAAGATATTG AATCAACATG GGCAGGAATT | 3600 |
| | | 3660 |
| | AGACCATTAA TTTACGAAGA AGGCAAAGAC CCTTCTGAAA TCTCTCGTAA GGATGAAATT TGGGAAGGTA AATCAGGTTT ATTAA GTATT | 3720 |
| 35 | TGGGAAGGTA AATCAGGTTT ATTAACTATT GCAGGTGGTA AATTAACAGG CTATCGTCAC | 3780 |
| | ATGGCTCAAG ACATTGTTGA TTTAGTATCT AAACGCTTGA AAAAAGACTA CGGTTTAACA | 3840 |
| | TTTAGTCCAT GTAATACAAA AGGTCTGGCA ATTTCAGGTG GCGATGTAGG TGGTAGCAAG | 3900 |
| 40 | AACTTTGATG CGTTTGTAGA GCAAAAAGTA GATGTAGCTA AAGGATTCGG CATTGATGAA | 3960 |
| | GATGTTGCAA GACGTTTAGC ATCTAAATAT GGTTCAAATG TTGATGAATT GTTCAACATT | 4020 |
| 15 | GCGCAAACAT CTCAATACCA TGATAGCAAG TTACCATTAG AAATTTATGT AGAACTTGTT | 4080 |
| 45 | TATAGTATTC AACAAGAAAT GGTATACAAA CCTAACGATT TCTTAGTTCG TCGTTCTGGT | 4140 |
| | AAAATGTATT TCAATATTAA AGATGTATTA GATTATAAAG ATGCTGTCAT CGATATTATG | 4200 |
| 50 | GCAGATATGC TTGATTACTC TECAGCTCAA ATTGAAGCAT ATACTGAAGA AGTTGAGCAA | 4260 |
| | GCAATTAAAG AAGCGCAACA TGGAAATAAT CAACCAGCAG TTAAAGAATA ALTAATTTGT | 4320 |
| | ACAATCATAA ACTGGTGTCC TGTTTTAAGG GCATCAGTTT TTTTATACGA GATACATTAG | 4380 |

| GTTATTAAAG | GTGTGAGATG | ATGACTGAAA | AACAATTTAA | ATTAACTGTA | CAAGATAATA | 4500 |
|------------|------------|------------|------------|------------|------------|------|
| CGAATATTGA | AGTTAAAGTG | AATTTTACAG | ATGTAGATTC | AAAAGGAATT | ATTCATATAT | 4560 |
| TTCATGGTAT | GGCTGAACAT | ATGGAACGTT | ACGATAAATT | AGCACATGCA | CTTTCAAAGC | 4620 |
| ATGGCTTCGA | TGTGATACGT | CATAATCATC | GAGGACATGG | TATTAATATT | GATGAATCAA | 4680 |
| CAAGAGGGCA | TTACGATGAT | ATGAAACGAG | TTATCGGTGA | TGCCTTTGAA | GTAGCGCAAA | 4740 |
| CAGTGAGAGG | CAATGTTGAT | AAACCATACA | TTATAATCGG | ACATTCAATG | GGATCCGTTA | 4800 |
| TAGCTAGATT | GTTTGTAGAA | ACATATCCGC | AATATGTTGA | TGGTCTAATT | TTAAGTGGTA | 4860 |
| CTGGTATGTA | TTCATTATGG | AAAGGTTTAC | CAACCGTTAA | AGTGTTACAA | CTGATTACAA | 4920 |
| AAATTTATGG | TGCTGAGAAA | CGAGTTGAAT | GGGTTAACCA | GTTAGTATCA | AATAGTTTTA | 4980 |
| TARARANDAT | ACGTCCATTA | CGTACACAAA | GTGATTGGAT | TTCTAGTAAT | CCAATTGAAG | 5040 |
| TAGATAaCTT | TATTAAAGAT | CCATATAGTG | GaTTTAATGT | GTCAAATCAA | TTATTATATC | 5100 |
| AAACAGCCTA | TTATATGCTA | CATACATCAC | AATTAAAAA | TATGAAAATG | TTAAATCATG | 5160 |
| CCATGCCTAT | ATTATTAGTT | TCAGGATATG | ACGATCCTTT | AGGTGATTAT | GGTAAAGGGA | 5220 |
| TTTTAAAATT | GGCGAATATA | TATAGAAACG | CTGGCATnAA | AAATGTTAAA | GTGAATCTTT | 5280 |
| ATCATCATAA | ACGTCATGAA | GTGTTATTTG | AAAAnGATCA | TGACNAAATT | TGGGAÁGACT | 5340 |
| TGTTTAAATG | GTTGAATCAA | TTTTATAAAA | AATAAAGAAA | GTGGAATTAA | ATATGAATAA | 5400 |
| AAATAAGCCT | TTTATTGTAG | TAATTGTGGG | GCCAACTGCT | TGCAG | | 5445 |

(2) INFORMATION FOR SEQ ID NO: 74:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2569 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

40

45

50

5

10

15

20

25

30

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

TGGCTTGAAC TACGCCAATA AGTCCCCCTA GTACAAGAAT GAATACCATG ATATCGACCG 60
CTTCTATCGT ACCTTCAACC ATGCTACTTG TTATTTGTTC TGGTCCAGCT GGATGTTGCT 120
TTAATCTTTC ATAAGTATTC GGAATTGATA CCGGCTTATT AATTGCACCT GATTTAAATT 180
GTTCAATCTT AATTTTAACC CCCATTTTGT CTAGTTCCTG TTGCGTACCC GGAACCTTTT 240
TCACTTGGTT ATGAGGGTTA ACTATCTTTA GTTCTTGGGA TGAAGGTTCG TAAGAAAGTT 300
TAGAATATGC ACCAGCAGGA ATAACCCATG TTGCTATAAC TGCAACAACC GTTAAAATGA 360

| | | - |
|-----------|--|----------------|
| | TAATTGTATT TTCCACGGTT TCATCTCCTT CGACATTTAA CCTAGCATTT CTACCTTA | |
| 5 | GATTITATAA ATATAAATTA AGAAAGTGCA CCCCGCATCA AAATAGAGGC ATTATTIT | |
| | GGGGGTGCAC ATAAATAATA AAAATCATGC ATTTGACATA TAGTAATTGA AAAGCGTT | IC 600 |
| | AATTCAATTA CTTTTTAATC ACAGTACCTA CTTTACCCTC TAAGGCAGCA TCTAATTC | AT 660 |
| 10 | TTAATGATGT TATAAGCACA CTTCCTTTTG GATTGTTTTC AATAAATGAT ATGGCTGC | F T 720 |
| | CAATTTTTGG TAACATACTT CCTTTTGCAA ATTGATTTTC GTCTATATAT CGTTTTAA | FT 780 |
| | CATCAACATT TGTTGTTTTC AAAGGCTGTT GGTTTTCAGT GTTAAAATTA ATATATACA | AT 840 |
| 15 | AATCAATTGC TGTTAAAATA ATCAATTGAT CGCATTGAAT ATTAGCACCC AACAACGC | AC 900 |
| | TIGTTTTATC TTTGTCTATA ACTGCATCAA TACCTTTAAA ACCATCATGT TGCTCTCTA | NA 960 |
| | TTACTGGTAT ACCTCCACCA CCAGCAGCAA TAACGAGTGT ATCATTTTTA ATAAGTGTT | T 1020 |
| 20 | TAATACTCTC TAATTCAATA ATAGAGATGG GTTGTGGTGA AGGAACAACG CGTCTATAT | C 1080 |
| | CTCTTCCAGC ATCTTCAACA AATATAAATC CTTTTTCTTT TTGAATTTGT TCAGCTTCT | T 1140 |
| | CTTTGTTGTA AAATAACCCA ATTGGTTTTG AAGGATTGTT AAATGCCGGA TCATTTTCA | T 1200 |
| <i>25</i> | CAACTTCAAC TTGTGTCACT AGTGTTACCA CTTGTTTATC CATTCCAATA GAATGCAAT | T 1260 |
| | CATTITGTAA GCTTTCTTGT AATTGATAGC CGATGTAAGC TTGACTCATT GCGCCACAT | T 1320 |
| 30 | CAGCAAATGG AAATGCCGGA CCTTGGTTAT GTTCTGCAGC ATAGTTAAGT CCCAAATTA | A 1380 |
| | TGCTTCCAAC CTGTGGTCCA TTACCATGAC TAATAACAAT CTCATGTCCT TTTGTnATT | A 1440 |
| | AYCCTACTAA TGATTLCGCA GTATTTTTAA CAAGCTCGAG TLGGTYCTTG aGGTGATTT | n 1500 |
| 35 | CCTAAAGCAT TACCACCTAA TGCTACTACT ATTTTCGCCA TCATATTCAC TTCCTTATA | T 1560 |
| | CATTTAAAAT TCACCCAATG TAGCAACCAT GACTGCTTTG ATTGTATGCA TTCTGTTCT | C 1620 |
| | AGCTTCTTGG AATACAACTG AAGCTTTACT TTCGAATACT TCATCTGTAA CTTCCATTT | C 1680 |
| 40 | TCGAATACCA TATTTTTCAA AAATTTGTTG ACCTATTTTC GTATCAGCAT TATGGAAAG | A 1740 |
| | TGGTAAGCAA TGCTCAAAAA TAACATTTGG ATTACCAGTT TTATCCATTA TTTCTTTAT | T 1800 |
| | TACTTGATAT GGTTTCAATA ATTCAAGTCG TTCTTTCCAT ACTTCATCAG GTTCACCCA | T 1860 |
| 15 | TGATACCCAA ACATCAGTGT AAATTACATC CGAACCTTTT ACaCCTTGGT CaATATCAT | C 1920 |
| | TGTGATTAAT ATGTTGCCAC CATTTTCAGC GGCAATATTT TTACAGCGAT TTAATAATT | C 19 80 |
| | ATCTGTTGGA TTTAATTCTT TTGGACAAAC TAAATGGAAG TTCATACCCA TAATGGCAG | C 2040 |
| 50 | ACCTTGCATT AATGCATTTG CAACGTTATT ACGACCATCT CCAACATATG TAAAGTTAA | T 2100 |
| | ATCTGCATAA TCTTTTTTTA AGACTTCTTT TGCTGTTAAG AAATCAGCAA GAACTTGAG | T 2160 |

| TTCTACTGTT | CTTTGTGAAA | AACCACGGTA | TTCAATGCCA | TCATACATTC | CACCAAGCAC | 2280 |
|-------------|-------------|-------------|------------|------------|------------|------|
| ACGTGCAGTA | TCTTTAGTTG | TTTCTTTTTT | ACCCATTTGT | GATCCAGTTG | GGCCTAAATA | 2340 |
| AGTTACATTT | GCACCTTGAT | CATGCGCTGC | AACTTCAAAT | GCACATCGCG | TTCTTGTAGA | 2400 |
| ATCTTTTTCA | AATAACAGTG | CAATATTTTT | ATTTTTTAAC | ATAGGCTTTT | CAGTGCCAAT | 2460 |
| ATATTTAGCA | CGTTTTAAAT | CCTCGGAGAG | TGTTAATAAG | GTTCTACCTC | TTGTCGTGAA | 2520 |
| AAGTCTAATA | AAGTTAAAAA | ACTTCTGTTT | CGTANATTTT | TCATTAAna | | 2569 |
| (2) INFORMA | TION FOR SE | Q ID NO: 75 | 5 : | | | |
| | | | | | | |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1273 base pairs(B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

20

25

30

35

4Ô

45

50

5

10

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

| CCTGGAACCA | TCCaATCGtG | CaAATCtTGa | AAGaGAATAC | GCAACAACAA | TTAAATGTAT | 60 |
|------------|------------|------------|------------|------------|------------|------|
| TGGAACACTA | TATTCCAAAT | GACCATCCAG | CACTCGTTGA | ATTAAAAATA | TGGGAACGTT | 120 |
| GGTTACATAA | ACAAGGTTAC | AAAGACATCC | ATTTAGATAT | TACTGCGCAC | CACCTAGATC | 180 |
| CTATTACACA | GGTTTATTTA | TTCAATGTCA | TTTTGCTGAA | AATGAATCTC | GAGTTTTAAC | 240 |
| AGGTGGTTAT | TACAAAGGAA | GCATCGAAGG | GTTTGGATTA | GGATTAACAC | TTTAAGTAAG | 300 |
| GGAGTATGCA | CAATGTTAAG | AATCGCCATA | GCCAAAGGAC | GTCTAATGGA | TAGTTTAATT | 360 |
| AACTATTTAG | ATGTAATTGA | ATATACGACA | TTATCAGAAA | CATTAAAAAA | TAGAGAACGC | 420 |
| CAATTATTAT | TAAGTGTAGA | TAATATTGAA | TGCATTTTAG | TAAAAGGAAG | TGACGTGCCA | 480 |
| ATCTĀTGTGG | AACAAGGAAT | GGCAGACATA | GGCATTGTTG | GTAGCGACAT | ATTAGATGAG | 540 |
| CGCCAATATA | ATGTTAATAA | TTTGTTGAAT | ATGCCTTTTG | GAGCATGTCA | TTTTGCGGTT | 600 |
| GCAGCGAAAC | CTGAAACGAC | CAATTATCGT | AAAATCGCAA | CGAGTTATGT | TCATACTGCT | 660 |
| GAAACATATT | TTAAATCAAA | AGGTATTGAT | GTCGAATTGA | TTAAATTGAA | TGGCTCTGTT | 720 |
| GAATTGGCCT | GTGTTGTAGA | TATGGTAGAC | GGAATTGTCG | ACATCGTTCA | AACAGGTACT | 780 |
| acgctaaaag | CGAACGGACT | GGTTGAAAAG | CAACATATTA | GTGATATCAA | TGCAAGATTA | 840 |
| ATAACTAATA | AAGCAGCTTA | TTTTAAAAAA | TCACAATTAA | TAGAGCAATT | TATTCGCTCT | 900 |
| TTGGAGGTGT | CTATTGCCAA | TGCTTAATGC | ACAACAATTT | TTAAATCAAT | TTTCATTAGA | 960 |
| AGCACCATTA | GATGAGTCAT | TGTATCCAAT | TATTCGCGAT | ATTTGTCAGG | AAGTTAAAGT | 1020 |

| TTTAGaAATT | AGTCATGAMC | AAATTAAAGC | AGCATTTGAC | ACATTAGATG | AAAAAACAAA | 1140 |
|------------|------------|------------|------------|------------|------------|------|
| ACAAGCATTA | CAACAAAGTT | ATGAAAGAAT | TANAGCATAT | CAaGAAaGTA | TtaAACAGaC | 1200 |
| Gaatcaacag | TTAGAAGAAT | CAGTGGaGTG | tTrTGaAATA | TACCATCCmC | taGaAAGTGT | 1260 |
| CGGTATTTAT | GTG | | | | | |
| | | | | | | 1273 |

(2) INFORMATION FOR SEQ ID NO: 76:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1308 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:

| | • | * | , | L | 70. | | |
|---|-------------|--------------|------------|--------------|------------|------------|------|
| | GTTGATAAAT | TAAAAATGTT | TTTATCAGAT | ` ATTCAAAGTT | ACCAACAATA | TAGTAAAGAT | 60 |
| | CATCCGGTGT | ATCAGTTAAT | TGATAAATTT | TATAATGATC | ATTATGTTAT | TCAATACTTT | 120 |
| | AGTGGACTTA | TTGGTGGACG | TGGACGACGT | GCAAATCTTT | ATGGTTTATI | TAATAAAGCT | 180 |
| | ATCGAGTTTG | AGAATTCAAG | TTTTAGAGGT | TTATATCAAT | TTATTCGTTT | TATCGATGAA | 240 |
| | TTGATTGAAA | GAGGCAAAGA | TTTTGGTGAG | GAAAATGTAG | TTGGTCCAAA | CGATAATGTC | 300 |
| | GTTAGAATGA | TGACAATTCA | TAGTAGTAAA | GGTCTAGAGT | TTCCATTTGT | CATTTATTCT | 360 |
| • | GGATTGTCAA | AAGATTTTAA | TAAACGTGAT | TTGAAACAAC | CAGTTATTTT | AAATCAGCAA | 420 |
| • | TTTGGTCTCG | GAATGGATTA | TTTTGATGTG | GATAAAGAAA | TGGCATTTCC | ATCTTTAGCT | 480 |
| • | rcggttgcat | ATAGAGCTGT | TGCCGATAAA | GAACTTGTGT | CAGAAGAAAT | GCGATTAGTC | 540 |
| | | | | | | GAAAAATGAT | 600 |
| | _ | | | | | TGTCAATGAA | 660 |
| | GATTAACTT | CACCAAATCC | GTTCCATCTT | ATTTATAGTA | TTTTATCTAA | ACATCAATCT | 720 |
| C | CCTCAATTC | CAGATGATTT | AAAATTTGAA | AAAGATATAG | CACAAATTGA | AGATAGTAGT | 780 |
| C | GTCCGAATG | TAAATATTTC | AATTGTGTAC | TTTGAAGATG | TGTCTACAGA | AACCATTITA | 840 |
| G | ATAATGATG | AATATCGTTC | GGTTAATCAA | TTAGAAACTA | TGCAAAATGG | TAATGAAGAT | 900 |
| G | TTAAAGCAC | AAATTAAACA | CCAACTTGAT | TATCGATATC | CATATGTAAA | TGATACTAAA | 960 |
| A | AGCCCTCAA | AACAATCTGT | TTCTGAATTG | AAAAGACAAT | ATGAAACAGA | AGAAAGTGGC | 1020 |
| A | CAAGTTACG . | AACGAGTAAG | GCAATATCGT | ATCGGTTTTT | CAACGTATGA | ACGACCTAAA | 1080 |
| T | TTCTAAGTG . | AACAAGGTAA . | ACGAAAAGCG | AATGAAATTG | GTACGTTAAT | GCATACAGTG | 1140 |

55

10

. 15

20

25

30

35

40

45

| | GATGGATTAA TCGATAAACA TATTATCGAA GCAGATGCGA AAAAAGATAT CCGTATGGAT | 1260 |
|----|--|------|
| | GAAATAATGA CATTTATCAA TAGTGATTAT ATTCGATATT GCTGAAGC | 1308 |
| 5 | (2) INFORMATION FOR SEQ ID NO: 77: | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1431 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 15 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 77: | |
| | GATGCCATTN ATNNGTATGC AAGAAGTTGT TCCGGGTTCA GGTGGATTaC CAGTTGGTAC | 60 |
| | TGGTGGTAAG ACGTTACTAA TGCTTTCAGG CGGTATAGAC TCACCAGTTG CTGGGATGGA | 120 |
| 20 | AGTGATGAGA CGTGGCGTAA CAATTGAAGC GATTCATTTC CATAGTCCAC CATTTACAAG | 180 |
| | TGATCAAGCA AAAGAAAAAG TTATTGAATT GACACGTATT TTAGCTGAAC GTGTTGGACC | 240 |
| | AATTAAATTG CATATTGTAC CATTTACAGA ATTGCAAAAA CAGGTAAATA AAGTTGTACA | 300 |
| 25 | TCCAAGATAT ACAATGACTT CAACGAGACG TATGATGATG CGTGTTGCTG ATAAATTAGT | 360 |
| | ACATCAAATA GGGGCTTTAG CTATTGTAAA TGGTGAAAAC CTAGGGCAGG TAGCCAGTCA | 420 |
| 30 | AACACTTCAT AGCATGTATG CAATTAATAA TGTAACTTCT ACTCCTGTAT TACGTCCTTT | 480 |
| οŲ | ATTAACTTAC GATAAAGAAG AAATTATTAT TAAATCGAAA GAAATTGGTA CATTTGAAAC | 540 |
| | ATCTATTCAA CCATTTGAAG ATTGTTGTAC AATTTTCACC CCTAAAAATC CAGTAACCGA | 600 |
| 35 | ACCAAACTTT GATAAGGTAG TCCAATATGA AAGTGTCTTT GATTTTGAAG AGATGATTAA | 660 |
| | TCGTGCTGTT GAAAATATTG AAACACTTGA AATAACTAGT GATTATAAAA CTATTAAAGA | 720 |
| | ACAGCAAACA AACCAATTAA TAAACGACTT TTTATAAATA AAATCCTAGA GTAAATTTAA | 780 |
| 40 | ACATAAGGGG ATGTTAAACT ATGGATTTGA ACTTAACGAT GATTATAATC ATAATTTTAT | 840 |
| | TTGGTTTTAT CGCGGCGTTT ATAGATTCGG TTGTAGGGGG TGGCGGTTTA ATTTCTACGC | 900 |
| | CAGCATTATT AGCAATCGGT CTACCACCAT CTGTGGCTTT AGGTACAAAT AAATTGGCAA | 960 |
| 45 | GTTCGTTTGG TTCTTTAACT AGTACGATAA AGTTTATAAG GTCCGGTAAA GTGGACTTAT | 1020 |
| | ATGTTGTTGC CAAATTATTT GGTTTTGTAT TTTTGGCATC TGCATGTGGC GCATATATTG | 1080 |
| 50 | CAACGATGGT TCCGTCACAA ATATTGAAAC CTTTAATCAT CATTGCACTT TCGTCGGTGT | 1140 |
| | TTATATTCAC ATTACTTAAA AAAGATTGGG GCAATACACG CACGTTTACT CAATTTACAT | 1200 |
| | TTAAGAAAGC CATAATATTT GCAGCACTTT TTATATTAAT CGGCTTTTAT GATGGATTTG | 1260 |

| | The Congress of the Congress o | - |
|------------|--|------|
| | TAAGTGCAGC AGGAAATGCT AAAGTTTTGA ACTTTGCTTC TAATATAGGT GCGCTTGTAT | 1380 |
| 5 | TATTTATGGT ATTAGGACAA GTAGATTATG TAATAGGTTT AATTATGGCT A | 1431 |
| - | (2) INFORMATION FOR SEQ ID NO: 78: | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 4403 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 15 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 78: | |
| | AATATTATTT TAAATTCAAT ATTTATTGGT GCATTTATTT TAAACTTATT ATTCGCCTTT | 60 |
| | ACCATTATTT TCATGGAAAG ACGTTCTGCC AATTCTATCT GGGCTTGGTT ACTAGTCTTA | 120 |
| 20 | GTTTTCTTGC CTTTATTCGG CTTCATTTTA TACTTACTAT TAGGACGACA AATTCAACGT | |
| | GACCAAATTT TCAAAATTGA TAAGGAAGAT AAAAAAGGAT TAGAGTTAAT CGTTGATGAG | 180 |
| | CAATTAGCTG CTTTAAAAAA TGAAAACTTT TCAAATTCCA ATTATCAAAT TGTAAAATTT | 240 |
| 25 | AAAGAAATGA TTCAAATGTT GTTATATAAT AACGCAGCAT TTTTAACAAC AGACAACGAT | 300 |
| | TTAYFFEAT ACACAGACGG CCAAGAAAA TTTGATGACC TAATACAAGA CATCCGTAAT | 360 |
| | GCTACTGATT ATATTCATTT TCAGTACTAT ATTATTCAAA ATGATGAATT AGGTCGTACC | 420 |
| 30 | ATTTTAAATG AACTTGGTAA AAAAGCGGAA CAAGGTGTAG AAGTTAAAAT TCTTTATGAT | 480 |
| | GACATGGGTT CTCGTGGACT GCGTAAAAAA GGCTTACGCC CGTTTCGCAA TAAAGGTGGA | 540 |
| 35 | | 600 |
| J 5 | CATGCTGAAG CATTTTCCC ATCAAAATTA CCTTTAATTA ACTTGCGTAT GAACAATCGA | 660 |
| | AACCATCGAA AAATTGTTGT AATAGATGGG CAAATTGGAT ATGTTGGTGG TTTTAATGTT | 720 |
| 40 | GGTGĀTGAGT ACTTAGGTAA ATCAAAAAAA TTCGGCTATT GGCGAGATAC GCATTTACGA | 780 |
| | ATTGTCGGGG ATGCAGTGAA TGCATTGCAA TTACGATTTA TTCTAGATTG GAATTCACAA | |
| | GCCACACGTG ACCACATCTC CTATGATGAT CGTTATTTCC CAGATGTAAA TTCTGGTGGA | 900 |
| 45 | ACAATTGGCG TTCAAATAGC TTCTAGTGGT CCTGACGAAG AATGGGAACA GATTAAATAC | 960 |
| | GGCTATTTGA AAATGATTTC ATCTGCTAAA AAATCGATTT ATATTCAATC TCCCTATTTC | 1020 |
| | ATACCTGATC AAGCCTTTTT AGATTCTATT AAAATTGCGG CATTAGGTGG TGTTGATGTC | 1080 |
| 50 | AATATCATGA TTCCTAATAA ACCTGACCAT CCGTTTGTTT TTTGGGCTAC TTTAAAAAAT | 1140 |
| | GCAGCATCCT TATTAGATGC CGGTGTTAAA GTATTTCACT ACGACAATGG CTTTTTACAC | 1200 |
| | TCAAAAACAC TTGTTATAGA TGATGAAATT GCAAGTGTGG GAACAGCTAA TATGGACCAT | 1260 |

| | AAATTAAAAC | AAGCTTTTAT | AGATGATTTA | GCAGTATCTT | CTGAATTAAC | AAAAGCACGT | 1380 |
|----|------------|------------|------------|------------|------------|------------|------|
| | TATGCTAAGC | GAAGTCTTTG | GATTAAATTT | AAAGAAGGTA | TTTCACAATT | ATTGTCACCT | 1440 |
| 5 | ATCTTATAAA | ATAGAAATAT | GAGGAGTGTA | aCTTTAATGC | AACAATCAGA | CGTCATTAGT | 1500 |
| | GCTGCCAAAA | AATATATGGA | ATCTATTCAT | CAAAATGATT | ATACAGGCCA | TGATATTGCG | 1560 |
| 10 | CATGTATATC | GTGTCACTGC | TTTAGCTAAA | TCAATCGCTG | AAAATGAAGG | TGTTAATGAT | 1620 |
| | ACTTTAGTCA | TTGAACTCGC | ATGTTTGCTT | CATGATACCG | TTGACGAAAA | AGTTGTAGAT | 1680 |
| | GCTAACAAAC | AATATGTTGA | ATTGAAGTCA | TTTTTATCTT | CTTTATCACT | ATCAACCGAA | 1740 |
| 15 | GATCAAGAGC | ACATTTTATT | TATTATTAAT | AATATGAGCT | ATCGCAATGG | CAAAAATGAT | 1800 |
| | CATGTCACTT | TATCTTTAGA | AGGTCAAATT | GTCAGGGATG | CAGATCGTCT | TGATGCTATA | 1860 |
| | GGCGCTATAG | GTGTTGCACG | AACATTTCAA | TTTGCAGGAC | ACTTTGGTGA | ACCTATGTGG | 1920 |
| 20 | ACAGAACATA | TGTCACTAGA | TAAGATTAAT | GATGATTTAG | TTGAACAGTT | GCCACCATCT | 1980 |
| | GCAATTAAAC | ATTTCTTTGA | AAAATTACTT | AAGTTAGAAT | CTTTAATGCA | TACAGATACG | 2040 |
| | GCGAAGATGA | TTGCTAAAGA | ACGTCACGAC | TTTATGATGA | TGTACTTGAA | ACAGTTTTTT | 2100 |
| 25 | ACGGAATGGA | ATTGTCACGA | CTAGACATTG | AAGTTGTAGT | ATGATGATGC | GATGTAATGG | 2160 |
| | CCTCTTCTTC | TGGAAGCTTG | GTGTCATGCC | ATGTTACTTT | GATGTGTTGT | TGTGGGAGCT | 2220 |
| 30 | TGGTGACATG | TCATGCTACT | TTGATGTGCT | GGTACCACGA | TGCGTCTTGA | TGTAGTGCTA | 2280 |
| 30 | TGATGTGGCA | TTGCGGTGTT | ATGGTGTTAT | AGACAGGTTT | GGCGTTGATG | CCATGTTACT | 2340 |
| | TTGATGTGCT | GGTACCACGA | TGCGACTTGA | TGTAGTGCTA | TGATGTGGCA | TTGCGGTGTT | 2400 |
| 35 | ATGGTGTTAT | AGACCGGTTT | GATGTTGATG | CCATGTTACT | TTGATGTGCT | GGTGCTACGA | 2460 |
| | TGCGACTTGA | TGTAGTGCTA | TGATGTGGCG | TTGCGCTGTT | ATGGTGTTAT | AGCCAGGTTT | 2520 |
| | GGTGTTGATG | TCATGCCGTT | ACGATTCTAT | GATATGTTGT | TGGGACGTTG | CAATGTGTAT | 2580 |
| 40 | TATGCCGTTG | TGACGTTATT | ATTTCACACT | GTTACATGTA | TAAGTGAATT | GCTGTGGAAA | 2640 |
| | TTTGCGACAT | ATACTGCTAC | ACTGATGAAT | CATTGTGTCA | AGATGACATT | GCGATGAAGA | 2700 |
| | ATGACAACTC | TGTTATTAAC | CACTTTTTAC | ATACTGAAAA | CTCGTTAATA | TTATTTCAAA | 2760 |
| 45 | TAAAAACAGC | AGTAGGATGA | CTTTCACATT | TGAAATCATC | TTACTGCTGT | TTCTATTTAT | 2820 |
| | CACATATTGT | ATAATGTGAC | ACTAAGTTTC | GCTATTGAAG | CGAAAAATAA | TGTGCGCCCT | 2880 |
| 50 | ATAAAGTTAA | AATTATCTTC | AACTTTTAGG | GTGCACATTA | TTTGGACTTG | CTAAGGTTAT | 2940 |
| 50 | TTCTTTTTCT | TTTTAGACAC | AACTTGTGTG | TTTTTGCCTT | TTTTATTGct | GCCGCCGTTG | 3000 |
| | TGCTCTCTTT | CATACGCTTC | AATGAAAGGT | TGTACTTCTT | TTTTAGCGAC | TTTTCATAA | 3060 |

| | CCAAGTGCT | G ATGCTGAGCT | TAATGAAAT | C CAGATAATC | A TAATTGGTGA | AATGACCATC | 3180 |
|----|------------|--------------|--------------|--------------|--------------|--|------|
| _ | ATCATGTAA | C CCATTTGACC | TTGTTCGTC | r ggcatcgtt | TACTTGATAC | ATATGCTTGG | 3240 |
| 5 | ATAAAGTAT | A AAACACCGGC | AATAATTGT/ | A ATCCAAATAT | CAGGACGTCC | TAAATCGAAC | 3300 |
| | CATAAGAAG | T GTGGATATTI | · AAACAAACCI | TCTACAAGTT | GGTCTTTAAG | TACAAAGTAT | 3360 |
| 10 | | | | | | ACTCTTAATC | 3420 |
| | | | | | | CATTTTTTCT | 3480 |
| | | | | | | CTTTGCAACT | 3540 |
| 15 | | | | | | TGGTAATAAA | 3600 |
| | | | | | | GTTTAATAAG | 3660 |
| | TTATTTCCCA | ACCAATCCAA | TACATTTTTC | ATTGGATCTA | CGAATGTATT | GTAGAAAAAy | 3720 |
| 20 | CWCtACGTTT | TTCAGGTTTA | GAATAGTCAC | AACCAGCCAA | AAAGACCATA | ATACCTAAAA | 3780 |
| | ATAATGGTAG | TAACGCTTTT | TTCTTCATTT | TTCCACCTCT | ATCATTATAT | TCACATAGGA | 3840 |
| 25 | | CACATTAATG | | | | | 3900 |
| 25 | | CTTTGAATCG | | | | | 3960 |
| | | AAATTCCATT | | | | | 4020 |
| 30 | | TAGCTACAAT | | | | | 4080 |
| ٠ | | CTGGTATATG | | | | | 4140 |
| | | CmACATGAGT | | | | | 4200 |
| 35 | | | | | | ATCATCTTGA | 4260 |
| | | | | | | ATTAACGATA | 4320 |
| | | CATATTGATG | | TGTAATTCTT | TAGCTAATAC | AAGTTTATCG | 4380 |
| 40 | TTTCCTTTTA | AAGCTGATTC | ACC | | | According to the control of the cont | 4403 |

(2) INFORMATION FOR SEQ ID NO: 79:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1808 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double (D) TOPOLOGY: linear

50

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

TGGANCCAAT ATTAGAAATG ATTAAAACAT TAACAGGTAT TAATAGTCCT TCAGGAGNCA 60

| | TAACAAATAA | AGGTGCGTTA | TTAATAACAG | TGCCAGGCAA | AAATGATGAA | GTACAACGCT | 180 |
|------------|------------|------------|------------|------------|------------|------------|------|
| _ | GTATTACTGC | TCATGTTGAT | ACTITAGGTG | CaATGGTTAA | AGAAATTAAA | GAAGATGGTC | 240 |
| 5 | GCTTaGCAAT | AGAATTAATT | GGAGGATTCA | CGTATAACGC | GATTGAGGGT | GAATATTGCC | 300 |
| | AAATTAAAAC | TGATGCTGGT | CAAATATATA | CAGGAACAAT | TTGTCTGCAT | GAAACAAGTG | 360 |
| 10 | TTCATGTATA | TAGAAATAAT | CATGAAATAC | CTAGAGATCA | AAAGCATATG | GAAATAAGAA | 420 |
| | TTGATGAAGT | AACTACATCA | GAAGAAGATA | CAAAGAGTTT | AGGTATTTCA | GTAGGTGATT | 480 |
| | TTGTTAGCTT | TGATCCACGT | ACAGTTATCA | CGTCATCAGG | TTTTATTAAA | TCTCGTCATT | 540 |
| 15 | TAGATGATAA | AGCTAGCGTA | CGgTtGATAC | TACAATTACT | AAAGAAATTA | AAAGAAGAGC | 600 |
| | TTATAATAAA | ACCACATACA | ACGCAATTTT | ATATTTCTAA | TAACGAAGAA | ATAGGTTACG | 660 |
| | GTGCAAATGC | ATCAATTGAT | TCGAAAATCA | AAGAATATAT | TGCATTAGAT | ATGGGCGCGT | 720 |
| 20 | TGGGAGACGG | TCAAGCATCG | GATGAATATA | CAGTTTCTAT | TTGTGCCAAA | GATGCTTCAG | 780 |
| | GTCCATATCA | TAAGCAATTG | AAATCGCACC | TAGTTAATCT | TTGCAAAATA | AATAACATTC | 840 |
| | CATATAAAGT | AGACATATAT | CCATATTATG | GTTCAGATGC | TTCAGCAGCT | TTACATGCTG | 900 |
| 25 | GTGCGGATAT | CAGACATGGT | TTATTTGGCG | CTGGCATTGA | ATCATCTCAT | GCAATGGAAC | 960 |
| | GAACACATAT | TGATTCTATT | AAAGCGACAG | AGAAATTACT | ATATGCATAT | TGCTTATCAC | 1020 |
| 30 | CAATTGAGTA | AACAATTAGT | GTTGACAAAT | GTGaACGACC | TATGTAATAT | AATGAACTAT | 1080 |
| | AAAAATAATT | AGAATTTTCT | AAAGAAATAG | TAGCAGATAT | GAAACGTAGC | AAATAGAAAG | 1140 |
| | CTAATGGGTG | ATGGGAATTA | GCACGCCATA | TCTTGTGAAT | TGGACTTTGG | AAAACAATTG | 1200 |
| 35 | AATGAGTTTT | GAAAGTGAAC | ATGAATTATG | TTAACTAAGG | TGGCACCACG | GTAACGCGTC | 1260 |
| | CTTACAGGTA | TATGCGTTAT | GTGGTGTCTT | TTTATTTAGA | CAAAATGTAG | TAGTTAATTA | 1320 |
| | AAGGTAGCAA | CAGAAAGTTA | GTGGATGATG | TGAACTAACA | CCGAGATTAA | TGAAATTGGG | 1380 |
| 40 | TTTTGTCTGC | AACAGAAAAA | TTATATATAG | TAAAGAGTGA | ACTATGAATA | TTTCGAATAT | 1440 |
| | TCGGTTAATT | TAGGTGGTAC | CACGCGTCAc | nTCCTTTATA | TTGATAAGGA | TGCTGGCGCT | 1500 |
| | TTTTTGAAAG | GAGCGTATAG | AATGGATATA | TTTTATAAAA | AAATAAAAGC | AAATGTAACG | 1560 |
| 45 | CCCGAAGTTT | TAGCACAACT | TCATTCCAAG | AAGaTCATTT | TGGAAAGTAC | AAATCAACAA | 1620 |
| | CAAACTAAAG | GTCGCTATTC | AGTTGTTATT | TTTGATATTT | ATGGCACTTT | AACTTTAGAT | 1680 |
| 50 | AATGATGTAT | TATCAGTAAG | TACTTTAAAA | GAATCGTATC | AAATCACTGA | AAGACCGTAC | 1740 |
| - - | CATTATTTAA | CGACTAAnAT | AAATGAAGAC | TACCATAATA | TTCCAAGATG | AGGCAACTTA | 1800 |
| | AGTCATTA | | | | | | 1808 |

| TE OUGUENCE CHARACTERISTICS | (i) | SEQUENCE | CHARACTERISTICS |
|-----------------------------|-----|----------|-----------------|
|-----------------------------|-----|----------|-----------------|

- (A) LENGTH: 1320 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

| | (X1) | SEQUENCE | DESCRIPTION: | SEQ | ID | NO: | 80: |
|--|------|----------|--------------|-----|----|-----|-----|
|--|------|----------|--------------|-----|----|-----|-----|

| | TGGTCGTCAA TTTCTTGATT ATATCTATAA TCCTCATTTT CAATATTAGA GTCTGTAGAA | 60 |
|----|---|-------|
| | TCATCGATAT TATTATCATT CGCATGACTA GAAGCAGAAT CATTATTTTT ATCATTGCTT | 120 |
| 15 | TCTTCTTTTT TGAAGTCTTT ATTTATCAAG TAAATTTCTT CATCAAAATC AGCTTGTTGA | 180 |
| | GATGTATCAT CTTTATTTTG ATTAGAAAAA TGTGTAGCCT TTGATCTTTT TCTTTGCCGT | 240 |
| | CTTTTCTTAG ATGTATTCCT CGTAAATAAT TCTAATTCAT CTTTATCTTC ATTTGATTCT | - 300 |
| 20 | TGTTGATCGT TCTTCGTTTT ATCATCCATC AATACTCACA CCCTTTAATA AGATGGTAAA | 360 |
| | TGGGCACGGA ATCTTTCAAT AAATTTCTCT CCACGCTCTT CAAAAGTACT ATATTGATCC | 420 |
| | CAACTCGCAC AAGCAGGTGA CAATAATACA ACATCATTTG GTTCTATAAT ATCTTGTACT | 480 |
| 25 | TTATCAACAG CGTCTTCGAC ATTGTTCGCT TCAATGACCG ATTTCCCTTG ACTATTACCT | 540 |
| | AGTTTAGCAA ACTTAGCTTT CGTTTGTCCG AATACAACCA TCGCGCGAAC ATTTTCCATA | 600 |
| | TAAGGAATGA GTTCGTCAAA TTCATTCCCT CGATCCAAAC CACCACATAA CCAAATGATT | 660 |
| 30 | GGTTGATTAA ATGAATTTAA GGCAAACTGT GTTGCTAGCG TGTTTGTTGC TTTGGAATCA | 720 |
| | TTATAATATT TATTAGTTCT ATTAGTACCA ACATATTGCA ATCTATGCTC TATTCCTGAA | 780 |
| 35 | AATGTAGTTA AACTATCAAT AATTGCLTTA ATAGGTACAC CAGCANAATA CAAGCAAGCA | 840 |
| | CAGCTGCTAA TATATTTCTA AATTATGTTC ACCAGGCAAT ACTAGALCTT CAGTGTTAAT | 900 |
| | AATACGAACA CCTTTATAAA CGATAAAACC ATCTTLAATA TAAATACCAT CArCTLCTTG | |
| 40 | TTGAGTTGAG AAATACAATG TCTTAGCTTT TAATTCTTCC GACTCTATCA CTTGTCTTTG | 960 |
| - | | 1020 |
| | ATGATAATTA CAAATCAAAT AATCCTCTTC CGTTTGATTT TTATATATTT GCTTTTTAGC | 1080 |
| | ATTITGATAG TITTCTAAAT TITCATGGTA ATCTAGATGC GCCGAATAAA TGTTAGTAAT | 1140 |
| 45 | TATAGCAATG TGTGGTTTAT ACTTTTCGAT TCCAAGTAAC TGGAATGACG ACAACTCTGT | 1200 |
| | AACTAAATAA TCTGTAGGCT TTACTTCTTG TGCTACTTTA GATGCAACAT AACCAATATT | 1260 |
| 50 | GCCGGATAAT CTTCCAGTTA AGCGACTTTT TTTAAACATA TCTCCAATTA GAGAAGTAAC | 1320 |
| 50 | (2) INFORMATION FOR SEC ID NO. 81. | |

(2) INFORMATION FOR SEQ ID NO: 81:

(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 4280 base pairs

55

5

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5

| | (xi) | SEQUENCE DE | SCRIPTION: | SEQ ID NO: | 81: | | |
|----|------------|----------------|------------|------------|------------|------------|------|
| | TTTACACCAA | TCAAAAAATC | GAACTGATAT | AAATAAGTAC | AAAGCTTATC | TATCAATCCG | 60 |
| 10 | ATTTAGTTAT | . Уууу уууу уу | AAGCCACAGT | AATGTGGCTT | TTTGTTATAT | TCAGTATCAA | 120 |
| | AATGGTATCA | ATAGCCATTT | TCGGAAGTCA | AGAATGGCTT | AACAACGCGG | TTTAAAGCTA | 180 |
| | TCCAATACTA | CCTTCCATTT | CGAACTTGAT | TAAACGGTTC | ATTTCGACCG | CGTATTCCAT | 240 |
| 15 | TGGAAGTTCT | TTTGTAAATG | GTTCGATGAA | TCCCATAACA | ATCATTTCTG | TCGCTTCTTC | 300 |
| | TTCAGAAATA | CCACGACTCA | TTAGATAGAA | TAATTGTTCT | TCAGAAACTT | TTGAAACCTT | 360 |
| | GGCTTCATGT | TCTAATGATA | TTTGATCGTT | GAATACTTCG | TTATATGGAA | TTGTATCTGA | 420 |
| 20 | TGTTGATTCG | TTATCTAAGA | TTAATGTATC | ACATTCAATA | TTTGAACGAG | CACCTTTTGC | 480 |
| | TTTACGTCCA | AAATGAACAA | TACCGCGATA | AATAACTTTA | CCACCATTTT | TAGAAATAGA | 540 |
| ?5 | TTTAGAAACA | ATTGTAGAAG | ATGTATTAGG | TGCTTTATGA | ATCATTTTAG | CACCGGCATC | 600 |
| .0 | TTGAACTTGT | CCTTTACCAG | CAAATGCAAT | AGATAATGTA | CTACCTTTTG | CACCTTCACC | 660 |
| | TAAAAGAACA | CAGTTTGGAT | ATTTCATCGT | TAACTTAGAA | CCTAAGTTAC | CATCTACCCA | 720 |
| 30 | TTCCATATTT | CCGTTTTCAT | AAACAAAAGT | ACGTTTTGTA | ACTAAATTGT | ATACATTGTT | 780 |
| | CGCCCAGTTT | TGAATCGTAG | TATAACGAAC | GTGCGCATCT | TTATGCACAA | TGATTTCCAC | 840 |
| | AACAGCAGAG | TGTAAAGAAC | TAGTTGTATA | AACTGGTGCA | GTACAACCTT | CTACGTAATG | 900 |
| 35 | TACAGAAGCA | CCTTCATCAG | CAATGATTAA | TGTACGTTCA | AATTGACCCA | TGTTCTCAGA | 960 |
| | GTTAATACGG | AAATAAGCTT | GTAGTGGCGT | ATCTAGTTTG | ATATTTTTAG | GTACATAAAT | 1020 |
| | GAAÇGAACCA | CCTGACCATA | CTGCTGAGTT | TAACGCCGCA | AATTTGTTAT | CTGCTGCAGG | 1080 |
| ‡O | TACTACAGAA | GCAAAGTATT | TTTTGAATAA | TTCTTCATTT | TCTTGTAAAG | CACTATCTGT | 1140 |
| | ATCTTTAAAG | ATAATACCTT | TTTCTTCAAG | TTCTTTTTCC | ATATTATGGT | AAACAACTTC | 1200 |
| | AGATTCATAT | TGAGCAGAAA | CACCAGCTAA | ATATTTTTGT | TCAGCTTCAG | GAATTCCTAA | 1260 |
| 15 | TTTATCGAAA | GTTCTTTTAA | TTTCTTCTGG | CACTTCATCC | CATGAACGTT | CAGCTTGTTC | 1320 |
| | TGAAGGCTTT | ACATAGTAAG | TAATGTCATC | GAAATTCAAT | TCTGATAAGT | CGCCACCCCA | 1380 |
| 50 | TTGAGGCATT | GGCATTTTAT | AAAACAATTT | TAATGATTTA | AGACGGAAAT | CTAACATCCA | 1440 |
| | TTCCGGCTCA | TTTTTCATGT | TAGAAATTTC | TCTAACGATA | TTCTCAGTTA | AACCACGTTC | 1500 |
| | TGATCTGAAA | ATGGACACAT | CATCGTCGTG | GAATCCATAT | TTATAATCCC | CAACATCAGG | 1560 |

| | TTTAATTCAT | GATGTAAACC | ATATTATAAC | AATGACATGA | CATCTTATAA | AAATTTTTAT | 1680 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | ACTTTTATAT | GTCTAATATC | AAAATTATCT | ATGATTAACA | GCATTCTATT | CTTCTTCAGT | 1740 |
| 5 | CGTACCTTCT | GCTTTACCTT | CTTTAGCAAC | AGTACCTTTT | TCCAATGCTT | TCCAAGCTAA | 1800 |
| | TGTGGCACAT | TTAATACGAG | CTGGGAATTG | AGATACACCT | TGCAATGCTT | CAATATCTCC | 1860 |
| | CATTTCTTCT | GTAATCACAT | AGTCTTCACC | AAGCATCATT | TTCGTAAATT | CTTGGCTCAT | 1920 |
| 10 | TTGCATTGCT | TCTCCAAGTG | AATGACCTTT | AACAGCTTGT | GTCATCATCG | ATGCACTTGC | 1980 |
| | CATTGAAATC | GAACAACCTT | CACCTTCAAA | CTTAGCATCT | TTTATAATGC | CGTCTTCTAT | 2040 |
| 15 | ATCAAATGTT | AGTCGTATAC | GGTCACCGCA | TGTCGGGTTA | TTCATATCTA | CTGTCATAGA | 2100 |
| | CCCGTTATCT | AATACACCTT | TATTTCTAGG | ATTTTTATAA | TGATCCATAA | TGACAGATCT | 2160 |
| | ATATAATTGA | TCTAGATTAT | TAAAATTCAT | AAGAGAAAA | CTCCTTCGTT | TGTTTCAAGG | 2,220 |
| 20 | CATTTATTAA | CTGATCAACG | TCTTCTTTCG | TGTTGTATAT | ATAAAAACTC | GCTCTAGCTG | 2280 |
| • | TTGAAGACAC | ATTTAACCAT | TTCATTAACG | GTTGCGCACA | ATGATGCCCA | GCTCTAACCG | 2340 |
| * | CTACACCTTC | TGTATCTACG | GCTGTAGCAA | CATCGTGTGG | ATGTACATCT | TGTAAATTAA | 2400 |
| 25 | ACGTTATTAC | ACCTGCACGA | CGATCCTTTG | GCGGGCCATA | AATTTCAATT | CCTTCAATTG | 2460 |
| | CAGACATTTG | CTCATAAGCA | TATATCGTTA | ATTCTTGTTC | ATATTTATGA | ATTGCATCAA | 2520 |
| | AACCTATGCG | TTCTAAATAG | CGAATAGCTT | CTGCAAGCCC | AATTGCTTGA | GCAATTAATG | 2580 |
| 30 | GAGTACCCGC | CTCAAATTTA | GTAGGTAAAT | CAGCCCATGT | TGCATCATAC | TTACTTACAA | 2640 |
| | AATCAATCAT | GTCGCCACCG | AACTCAATCG | GTTCCATTTT | TTGTAGTAAC | TCACGTTTAC | 2700 |
| 25 | CAAATAATAC | GCCAATACCT | GTTGGTCCAA | GCATTTTATG | ACCACTAAAA | CTATAAAAAT | 2760 |
| 35 | CAGCATTCAT | TTCTTGCATA | TCAAGTTTCA | TATGTGGTGC | TGctTGCGCC | CCATCAACAC | 2820 |
| | TGATAATTGC | ACCATGTTGA | TGAGCTATTT | CTGCAATGGT | TTTAACATCA | TTAATTGTAC | 2880 |
| 40 | CGAGCACATT | AGATATATGT | GCAATAGCAA | CGATCTTTGT | TTTATCATTA | ATCGTTTGCT | 2940 |
| | TAATATCCTC | GATGTTTAAT | TCACCGTCAG | CTGTCATTGG | TATAAATTTC | AATGTCGCAT | 3000 |
| | TTTTACGCTT | TGCTAACTGT | TGCCAAGGAA | CAATATTGGC | ATGATGTTCC | ATTTCAGTGA | 3060 |
| 45 | CAACAATTTC | ATCGCCCTCT | TCAACATTTG | CATCACCATA | GCTATGTGCT | ACAAGGTTAA | 3120 |
| | TCGACGCAGT | TGTTCCGCGT | GTAAAAATGA | TTTCTTCAAA | ATACTTCGCA | TTAATAAAAC | 3180 |
| | GACGAACGGT | TTCACGGGCA | TTTTCATAAC | CATCAGTTGC | CAATGATCCT | AATGTATGAA | 3240 |
| 50 | CACCACGATG | AACGTTTGAA | TTATAACGCT | TGTAGTAATC | TTCTAAAACA | TTTAACACTT | 3300 |
| | GCACAGGCGT | TTGACTTGTC | GCTGTTGAAT | CAAGATATGC | TAAACGTTTG | CCATTGACTT | 3360 |

| | CTTCATTCAC | GACCTTTCTT | AAATAAAAT | CCTAATCATT | TAAATACTGA | CGTTGTATTA | 3480 |
|----------|------------|------------|------------|------------|------------|------------|------|
| | GTCTTATACC | AATATCGACA | GTCTATATCT | ATTACAAACT | TTTATTTTCA | AAATATTATT | 3540 |
| 5 | TAGAAACTTT | GCGTTCAATT | ACTTCTCTCA | ATTGACGTTT | AACGTCTTCG | ATAGGTAATT | 3600 |
| | CACGTACTAC | TGGATCTAAG | AAACCATGTA | TAACAAGACG | TTCCGCTTCT | CTTTGAGAAA | 3660 |
| o | TACCACGACT | CATTAAATAG | TAAAGTTGAT | CTGGATCAAC | ACGACCTACT | GATGCAGCAT | 3720 |
| | GACCAGCTTG | TACATCATCT | TCATCAATTA | ATAAAATAGG | ATTCGCGTCA | CCACGAGCAT | 3780 |
| | GTTCAGATAA | CATTAATACA | CGTGATTCCT | GATTAGCAAT | TGATTTAGTT | CCACCATGCT | 3840 |
| 5 | TAATGTAGCC | GATACCATTA | AATACAGACG | ATGCATGTTC | TTTCATAACA | CCATGTTTAA | 3900 |
| | GGATATAACC | ATCTGTTTCT | TTACCATATT | GTACGATTTT | AGATGTTAGA | TTAATTTTTT | 3960 |
| | GTTCGCCTGT | ACCTACAACT | ACTGATTTAA | GTGAACTTGT | TGAACGATCA | CCAAATAAAT | 4020 |
| 0 | TTGTTGTATT | ATCAATAATT | TGGCTACCCT | CATTCATTAA | ACCTAGTGCC | CAATTAATTG | 4080 |
| | AGGCATCCGC | TTCAGTAATA | CCACGTCGAA | TGATATGACC | TGTAAAGCCT | TTATCCATAT | 4140 |
| | AGTCCACTGA | GCCATATGTG | ATATTTGAAT | TTGCACCAGC | AATCACTTCA | GAAATAATAT | 4200 |
| 5 | TtAATTGATT | TCCTTCACCA | GATGCATTTG | mTAAGTAATT | TTCAACATAT | GTGACTTCGG | 4260 |
| | CGCTTTCTTC | AGTAACGATG | | | | | 4280 |
| | | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 82:

30 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15598 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

35

40

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 82:

| TCnGACTCGA | ACGGTGMAAC | TALLCCGTTG | TaATTCCgGA | GgAAsCAAGG | TATGCCCATC | 60 |
|------------|---------------|--------------|------------|------------|------------|-----|
| TGCaAAGAAA | gaATGsAATG | AACTTTTTGG | AAATGTAGAA | GTGGTAAATA | AAGATAAAGG | 120 |
| ATATTACATT | CTGAGAAGTA | TAAAAGCTTG | AAATGAAATG | GATATTCTGT | TATAGTTATA | 180 |
| TAATGTAAAA | ATTTATGTTC | AATAAGTGTG | TACTTTTACG | TTAAATAGAT | AAGTTAATTA | 240 |
| AGAATAAATA | TAGAATCGAA | AATGGTGTCA | TCATTAGTGT | TGCCGTTTTC | TTTTTGTCTT | 300 |
| TTTATTAATA | TGCTTATGGT | ATTTAGCTAA | AAGCGGATCA | CATAATTTTT | GAGGGGTGAA | 360 |
| TCTGTTTGGC | AGGTCAAGTT | GTCCAATATG | GAAGACATCG | TAAACGTAGA | AACTACGCGA | 420 |
| | 1.00010001011 | mm> cc> > cm | | | | |

55

| | CIGGTAATT | r GTCATTAGAG | TTTGTGGATI | ACCGTTTAGG | AGAACCAAAA | TATGATTTAG | 60 |
|----|------------|--------------|------------|------------|------------|------------|------|
| | AAGAATCTAA | AAACCGTGAC | GCTACTTATG | CTGCACCTCT | TCGTGTAAAA | GTGCGTCTAA | 66 |
| 5 | TCATTAAAGA | AACAGGAGAA | GTTAAAGAAC | AAGAAGTCTT | TATGGGTGAT | TTCCCATTAA | 72 |
| | TGACTGATAC | AGGTACGTTC | GTTATCAATG | GTGCAGAACG | TGTAATCGTA | TCTCAATTAG | 78 |
| | TTCGTTCACC | ATCCGTTTAT | TTCAATGAAA | AAATCGACAA | AAATGGTCGT | GAAAACTATG | 84 |
| 10 | ATGCAACAAT | TATTCCAAAC | CGTGGTGCAT | GGTTAGAATA | TGAAACAGAT | GCTAAAGATG | 900 |
| | TTGTATACGT | ACGTATTGAT | AGAACACGTA | AACTACCATT | AACAGTATTG | TTACGTGCAT | 960 |
| 15 | TAGGTTTCTC | AAGCGACCAA | GAAATTGTTG | ACCTTTTAGG | TGACAATGAA | TATTTACGTA | 1020 |
| | ATACTTTAGA | GAAAGACGGC | ACTGAAAACA | CTGAACAAGC | GTTATTAGAA | ATCTATGAAC | 1080 |
| | GTTTACGTCC | AGGTGAACCA | CCAACTGTTG | AAAATGCTAA | AAGTCTATTG | TATTCACGTT | 1140 |
| 20 | TCTTTGATCC | AAAACGCTAT | GACTTAGCAA | GCGTGGGTCG | TTATAAAACA | AACAAAAAAT | 1200 |
| | TACATTTAAA | ACATCGTTTA | TTTAATCAAA | AATTAGCTGA | GCCAATTGTA | AATACTGAAA | 1260 |
| | CTGGTGAAAT | TGTAGTTGAA | GAAGGTACAG | TGCTTGATCG | TCGTAAAATC | GACGAAATCA | 1320 |
| ?5 | TGGATGTACT | TGAATCAAAT | GCAAACAGCG | AAGTGTTTGA | ATTGCATGGT | AGCGTTATAG | 1380 |
| | ACGAGCCAGT | AGAAATTCAA | TCAATTAAAG | TATATGTTCC | TAACGATGAT | GAAGGTCGTA | 1440 |
| | CGACAACTGT | AATTGGTAAT | GCTTTCCCTG | ACTCAGAAGT | TAAATGCATT | ACACCAGCAG | 1500 |
| 30 | ATATCATTGC | TTCAATGAGT | TACTTCTTTA | ACTTATTAAG | CGGTATTGGA | TATACAGATG | 1560 |
| | ATATTGACCA | TTTAGGTAAC | CGTCGTTTAC | GTTCTGTAGG | TGAATTACTA | CAAAACCAAT | 1620 |
| | TCCGTATCGG | TTTATCAAGA | ATGGAAAGAG | TTGTACGTGA | AAGAATGTCA | ATTCAAGATA | 1680 |
| 35 | CTGAGTCTAT | CACACCTCAA | CAATTAATTA | ATATTCGACC | TGTTATTGCA | TCTATTAAAĞ | 1740 |
| | AATICTTTGG | TAGCTCTCAA | TTATCACAAT | TCATGGACCA | AGCAAACCCA | TTAGCTGAGT | 1800 |
| 10 | TAACGCATAA | ACGTCGTCTA | TCAGCATTAG | GACCTGGTGG | TTTAACACGT | GAACGTGCTC | 1860 |
| | AAATGGAAGT | ACGTGACGTT | CACTACTCTC | ACTATGGCCG | TATGTGTCCA | ATTGAAACAC | 1920 |
| | CTGAGGGACC | AAACATTGGA | TTGATTAACT | CATTATCAAG | TTATGCACGT | GTAAATGAAT | 1980 |
| 15 | TCGGCTTTAT | TGAAACACCA | TATCGTAAAG | TTGATTTAGA | TACACATGCT | ATCACTGATC | 2040 |
| | AAATTGACTA | TTTAACAGCT | GACGAAGAAG | ATAGCTATGT | TGTAGCACAA | GCAAACTCTA | 2100 |
| | AATTAGATGA | AAATGGTCGT | TTCATGGATG | ATGAAGTTGT | ATGTCGTTTC | CGTGGTAACA | 2160 |
| 60 | | | , | ATATGGATGT | | | 2220 |
| | • | | | AAAATGATGA | | | 2280 |

| | CAGGTATGGA | ACACGTTGCA | GCACGTGATT | CTGGTGCGGC | TATTACAGCT | AAGCACAGAG | 240 |
|----|------------|------------|------------|------------|------------|------------|------|
| _ | GTCGTGTTGA | ACATGTTGAA | TCTAATGAAA | TTCTTGTTCG | TCGTCTAGTT | GAAGAGAACG | 246 |
| 5 | GCGTTGAGCA | TGAAGGTGAA | TTAGATCGCT | ATCCATTAGC | TAAATTTAAA | CGTTCAAACT | 2520 |
| | CAGGTACATG | TTACAACCAA | CGTCCAATCG | TTGCAGTTGG | AGATGTTGTT | GAGTATAACG | 2586 |
| 10 | AGATTTTAGC | AGATGGACCA | TCTATGGAAT | TAGGAGAAAT | GGCATTAGGT | AGAAACGTAG | 2640 |
| | TAGTTGGTTT | CATGACTTGG | GACGGTTACA | ACTATGAGGA | TGCCGTTATC | ATGAGTGAAA | 2700 |
| | GACTTGTGAA | AGATGACGTG | TATACTTCTA | TTCATATTGA | AGAGTATGAA | TCAGAAGCAC | 2760 |
| 15 | GTGATACTAA | GTTAGGACCT | GAAGAAATCA | CAAGAGATAT | TCCTAATGTT | TCTGAAAGTG | 2820 |
| | CACTTAAGAA | CTTAGACGAT | CGTGGTATCG | TTTATATTGG | TGCAGAAGTA | AAAGATGGAG | 2880 |
| | ATATTTTAGT | TGGTAAAGTA | ACGCCTAAAG | GTGTAACTGA | GTTAACTGCC | GAAGAAAGAT | 2940 |
| 20 | TGTTACATGC | AATCTTTGGT | GAAAAAGCAC | GTGAAGTTAG | AGATACTTCA | TTACGTGTAC | 3000 |
| | CTCACGGCGC | TGGCGGTATC | GTTCTTGATG | TAAAAGTATT | CAATCGTGAA | GAAGGCGACG | 3060 |
| | ATACATTATC | ACCTGGTGTA | AACCAATTAG | TACGTGTATA | TATCGTTCAA | AAACGTAAAA | 3120 |
| ?5 | TTCATGTTGG | TGATAAGATG | TGTGGTCGAC | ATGGTAACAA | AGGTGTCATT | TCTAAGATTG | 3180 |
| | TTCCTGAAGA | AGATATGCCT | TACTTACCAG | ATGGACGTCC | GATCGATATC | ATGTTAAATC | 3240 |
| | CTCTTGGTGT | ACCATCTCGT | ATGAACATCG | GACAAGTATT | AGAGCTACAC | TTAGGTATGG | 3300 |
| 30 | CTGCTAAAAA | TCTTGGTATT | CACGTTGCAT | CACCAGTATT | TGACGGTGCA | AACGATGACG | 3360 |
| | ATGTATGGTC | AACAATTGAA | GAAGCTGGTA | TGGCTCGTGA | TGGTAAAACT | GTACTTTATG | 3420 |
| 35 | ATGGACGTAC | AGGTGAACCA | TTCGATAACC | GTATTTCAGT | AGGTGTAATG | TACATGTTGA | 3480 |
| | AACTTGCGCA | CATGGTTGAT | GATAAATTAC | ATGCGCGTTC | AACAGGACCA | TATTCACTTG | 3540 |
| | tTACACAACA | ACCACTTGGC | GGTAAAGCGC | AATTCGGTGG | ACAACGTTTT | GGTGAGATGG | 3600 |
| 10 | AGGTATGGGC | ACTTGAAGCA | TATGGTGCTG | CATACACATT | ACAAGAAATC | TTAACTTACA | 3660 |
| | AATCCGATGA | TACAGTAGGA | CGTGTGAAAA | CATACGAGGC | TATTGTTAAA | GGTGAAAACA | 3720 |
| | TCTCTAGACC | AAGTGTTCCA | GAATCATTCC | GAGTATTGAT | GAAAGAATTA | CAAAGTTTAG | 3780 |
| 15 | GTTTAGATGT | AAAAGTTATG | GATGAGCAAG | ATAATGAAAT | CGAAATGACA | GACGTTGATG | 3840 |
| | ACGATGATGT | TGTAGAACGC | AAAGTAGATT | TACAACAAAA | TGATGCTCCT | GAAACACAAA | 3900 |
| | AAGAAGTTAC | TGATTAATAC | GCAATTTACA | AAACAGGCAA | AAAGATACTA | AGCTGAATTT | 3960 |
| 50 | TATTGATGAT | TCAGTTTAGT | ACTTTAAGCC | ATTTTAAATA | AATGCAAATC | AATCAAATAG | 4020 |
| | | | | | | | |

| | AAACCTGAAA | CAATCAACTA | CCGTACATTA | AAACCTGAAA | AAGATGGTCT | ATTCTGTGAA | 420 |
|------------|------------|--------------|------------|-------------------|----------------------|------------|------|
| | AGAATTTTCG | GACCTACAAA | AGACTGGGAA | TGTAGTTGTG | GTAAATACAA | ACGTGTTCGC | 426 |
| 5 | TACAAAGGCA | TGGTCTGTGA | CAGATGTGGA | GTTGAAGTAA | СТАААТСТАА | AGTACGTCGT | 432 |
| | GAAAGAATGG | GTCACATTGA | ACTTGCTGCT | CCAGTTTCTC | ACATTTGGTA | TTTCAAAGGT | 4380 |
| 10 | ATACCAAGTC | GTATGGGATT | ATTACTTGAC | ATGTCACCAA | GAGCATTAGA | AGAAGTTATT | 444(|
| | TACTTTGCTT | CTTATGTTGT | TGTAGATCCA | GGTCCAACTG | GTTTAGAAAA | GAAAACTTTA | 4500 |
| | TTATCTGAAG | CTGAATTCAG | AGATTATTAT | GATAAATACC | CAGGTCAATT | CGTTGCAAAA | 4560 |
| 15 | ATGGGTGCAG | AAGGTATTAA | AGATTTACTT | GAAGAGATTG | ATCTTGACGA | AGAACTTAAA | 4620 |
| | TTGTTACGCG | ATGAGTTGGA | ATCAGCTACT | GGTCAAAGAC | TTACTCGTGC | AATTAAACGT | 4680 |
| | TTAGAAGTTG | TTGAATCATT | CCGTAATTCA | GGTAACAAAC | CTTCATGGAT | GATTTTAGAT | 4740 |
| 20 | GTACTTCCAA | TCATCCCACC | AGAAATTCGT | CCAATGGTTC | AATTAGATGG | TGGACGATTT | 4800 |
| | GCAACAAGTG | ACTTAAACGA | CTTATACCGT | CGTGTAATTA | ATCGAAATAA | TCGTTTGAAA | 4860 |
| | CGTTTATTAG | ATTTAGGTGC | ACCTGGTATC | ATCGTTCAAA | ACGAAAAACG | TATGTTACAA | 4920 |
| 25 | GAAGCCGTTG | ACGCTTTAAT | TGATAATGGT | CGTCGTGGTC | GTCCAGTTAC | TGGCCCAGGT | 4980 |
| | AACCGTCCAT | TAAAATCTTT | ATCTCATATG | TTAAAAGGTA | AACAAGGTCG | TTTCCGTCAA | 5040 |
| | AACTTACTTG | GTAAACGTGT | TGACTATTCA | GGACGTTCAG | TTATTGCAGT | AGGTCCAAGC | 5100 |
| 3 <u>0</u> | TTGAAAATGT | ACCAATGTGG | TTTACCAAAA | GAAATGGCAC | TTGAACTATT | TAAACCATTC | 5160 |
| | GTAATGAAAG | AATTAGTTCA | ACGTGAAATT | GCAACTAACA | TTAAAAATGC | GAAGAGTAAA | 5220 |
| 35 | ATCGAACGTA | TGGATGATGA | AGTTTGGGAC | GTATTGGAAG | AAGTAATTAG | AGAACATCCT | 5280 |
| | GTATTACTTA | ACCGTGCACC | AACACTTCAT | AGACTTGGTA | TTCAAGCATT | TGAACCAACT | 5340 |
| | TTAGTTGAAG | GTCGTGCGAT | TCGTCTACAT | CCACTTGTAA | CAACAGCTTA | TAACGCTGAC | 5400 |
| 10 | TTTGACGGTG | ACCAAATGGC . | GGTTCACGTT | CCTTTATCAA | AAGAGGCACA | AGCTGAAGCA | 5460 |
| | AGAATGTTGA | TGTTAGCAGC | ACAAAACATC | TTGAACCCTA | AAGATGGTAA | ACCTGTAGTT | 5520 |
| | ACACCATCAC | AAGATATGGT | ACTTGGTAAC | TATTACCTTA | CTTTAGAAAG | AAAAGATGCA | 5580 |
| 15 | GTAAATACAG | GCGCAATCTT | TAATAATACA | AATGAAGTAT | TAAAAGCATA | TGCAAATGGC | 5640 |
| | TTTGTACATT | TACACACTAG | AATTGGTGTA | CATGCAAGTT | CGTTCAATAA | TCCAACATTT | 5700 |
| | ACTGAAGAAC | AAAACAAAAA | GATTCTTGCT | ACGTCAGTAG | GTAAAATTAT | ATTCAATGAA | 5760 |
| 50 | ATCATTCCAG | ATTCATTTGC | TTATATTAAT | GAACCTACGC | AAGAAAACTT | AGAAAGAAAG | 5820 |
| | ACACCAAACA | САТАТПТСАТ | CCATCCTACA | y Catataly Counce | 3.3.C.C.T.C.C.3.C.T. | | |

| | GAAGTATTCA | ACAGATTTAG | CATCACTGAT | ACATCAATGA | TGTTAGACCG | TATGAAAGAC | 600 |
|----|--------------|------------|------------|------------|------------|------------|------|
| _ | TTAGGATTCA | AATTCTCATC | TAAAGCTGGT | ATTACAGTAG | GTGTTGCTGA | TATCGTAGTA | 606 |
| 5 | TTACCTGATA | AGCAACAAAT | ACTTGATGAG | CATGAAAAAT | TAGTCGACAG | AATTACAAAA | 612 |
| | CAATTCAACC | GTGGTTTAAT | CACTGAAGAA | GAAAGATATA | ATGCAGTTGT | TGAAATTTGG | 618 |
| 10 | ACAGATGCAA | AAGATCAAAT | TCAAGGTGAA | TTGATGCAAT | CACTTGATAA | AACTAACCCA | 624 |
| | ATCTTCATGA | TGAGTGATTC | AGGTGCCCGT | GGTAACGCAT | CTAACTTTAC | ACAGTTAGCA | 6300 |
| | GGTATGCGTG | GATTGATGGC | CGCACCATCT | GGTAAGATTA | TCGAATTACC | AATCACATCT | 6360 |
| 15 | TCATTCCGTG | AAGGTTTAAC | AGTACTTGAA | TACTTCATCT | CAACTCACGG | TGCACGTAAA | 6420 |
| | GGTCTTGCCG | ATACAGCACT | TAAAACAGCT | GACTCAGGAT | ATCTTACTCG | TCGTCTTGTT | 6480 |
| | GACGTGGCAC | AAGATGTTAT | TGTTCGTGAA | GAAGACTGTG | GTACTGATAG | AGGTTTATTA | 6540 |
| 20 | GTTTCTGATA | TTAAAGAAGG | TACAGAAATG | ATTGAACCAT | TTATCGAACG | TATTGAAGGT | 6600 |
| | CGTTATTCTA | AAGAAACAAT | TCGTCATCCT | GAAACTGATG | AAATAATCAT | TCGTCCTGAT | 6660 |
| | GAATTAATTA | CACCTGAAAT | TGCTAAGAAA | ATTACAGATG | CTGGTATTGA | ACAAATGTAT | 6720 |
| 25 | ATTCGCTCAG | CATTTACTTG | TAACGCACGA | CATGGTGTTT | GTGAAAAATG | TTACGGTAAA | 6780 |
| | AACCTTGCTA | CTGGTGAAAA | AGTTGAAGTT | GGTGAAGCAG | TTGGTACAAT | TGCAGCCCAA | 6840 |
| | TCTATCGGTG | AACCAGGTAC | ACAGCTTACA | ATGCGTACAT | TCCATACAGG | TGGGGTAGCA | 6900 |
| 30 | GGTAGCGATA | TCACACAAGG | TCTTCCTCGT | ATTCAAGAGA | TTTTCGAAGC | ACGTAACCCT | 6960 |
| | AAAGGTCAAG | CGGTAATTAC | GGAAATCGAA | GGTGTCGTAG | AAGATATTAA | ATTAGCAAAA | 7020 |
| 35 | GATAGACAAC | AAGAAATTGT | TGTTAAAGGT | GCTAATGAAA | CAAGATCATA | CCTTGCTTCA | 7080 |
| | GGTACTTCAA | GAATTATTGT | AGAAATCGGT | CAACCAGTTC | AACGTGGTGA | AGTATTAACT | 7140 |
| | GAAĢGTTCTA | TTGAACCTAA | GAATTACTTA | TCTGTTGCTG | GATTAAACGC | GACTGAAAGC | 7200 |
| 40 | TACTTATTAA | AAGAAGTACA | AAAAGTTTAC | CGTATGCAAG | GTGTAGAAAT | CGACGATAAA | 7260 |
| | CACGTTGAGG | TTATGGTTCG | ACAAATGTTA | CGTAAAGTTA | GAATTATCGA | AGCAGGTGAT | 7320 |
| | ACGAAGTTAT | TACCAGGTTC | ATTAGTTGAT | ATTCATAACT | TTACAGATGC | AAATAGAGAA | 7380 |
| 45 | GCATTTAAAC | ACCGTAAGCG | TCCTGCAACA | GCTAAACCAG | TATTACTTGG | TATTACTAAA | 7440 |
| | GCATCACTTG . | AAACAGAAAG | TTTCTTATCT | GCAGCATCAT | TCCAAGAAAC | AACAAGAGTT | 7500 |
| | CTTACAGATG | CAGCAATTAA | AGGTAAGCGT | GATGACTTAT | TAGGTCTTAA | AGAAAACGTA | 7560 |
| 50 | ATTATTGGTA | AGTTAATTCC | AGCTGGTACT | GGTATGAGAC | GTTATAGCGA | CGTAAAATAC | 7620 |
| | GAAAAAACAG | CTAAACCAGT | TGCAGAAGTT | GAATCTCAAA | CTCAACTAAC | CCAAMAACAA | 3680 |

| | ATGTTGACGA | ATTCTCTTGT | TCAATGTTAA | TATATTAAAG | GTTGATGCAA | GCAGAACTTT | 780 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | GGAGGATAAA | TTATTGTCTA | AGGAAAAAGT | tGCACGCTTT | AACAAACAAC | ATTTTGTAGT | 786 |
| 5 · | TGGTCTTAAA | GAAACGCTTA | AAGCGTTAAA | GAAAGATCAA | GTTACATCTT | TGATTATTGC | 7920 |
| | TGAAGACGTT | GAAGTATATT | TAATGACTCG | CGTGTTAAGC | CAAATCAATC | AGAAAAATAT | 7980 |
| 10 | ACCTGTATCT | TTTTTCAAAA | GCAAACATGC | TTTGGGTAAA | CATGTAGGTA | TTAACGTCAA | 8040 |
| ,0 | TGCGACAATA | GTAGCATTGA | TTAAATGAGA | ATTAGTAAGT | GTTTTACTTA | CTAAATTTTA | 8100 |
| | TTTAACCTAA | AAATGAACCA | CCTGGATGTG | TGGGATTAAA | AAGTGAAGAG | AGGAGGACAT | 8160 |
| 15 | ATCACATGCC | AACTATTAAC | CAATTAGTAC | GTAAACCAAG | ACAAAGCAAA | ATCAAAAAAT | 8220 |
| | CAGATTCTCC | AGCTTTAAAT | AAAGGTTTCA | ACAGTAAAAA | GAAAAAATTT | ACTGACTTAA | 8280 |
| | ACTCACCACA | AAAACGTGGT | GTATGTACTC | GTGTAGGTAC | AATGACACCT | AAAAAACCTA | 8340 |
| 20 | ACTCAGCGTT | ACGTAAATAT | GCACGTGTGc | gTtTATCAAA | CAACATCGAA | ATTAACGCAT | 8400 |
| | ACATCCCTGG | TATCGGACAT | AACTTACAAG | AACACAGTGT | TGTACTTGTA | CGTGGTGGAC | 8460 |
| | GTGTAAAAGA | CTTACCAGGT | GTGCGTTACC | ATATTGTACG | TGGAGCACTT | GATACTTCAG | 8520 |
| 25 | GTGTTGACGG | ACGTAGACAA | GGTCGTTCAT | TATACGGAAC | TAAGAAACCT | AAAAACTAAG | 8580 |
| | AATTTAGTTT | TTAATTAAAT | CTTAAACTTA | AAATATTTAA | TATAAGGAAG | GGAGGATTTA | 8640 |
| | CATTATGCCT | CGTAAAGGAT | CAGTACCTAA | AAGAGACGTA | TTACCAGATC | CAATTCATAA | 8700 |
| 30 | CTCTAAGTTA | GTAACTAAAT | TAATTAACAA | AATTATGTTA | GATGGTAAAC | GTGGAACAGC | 8760 |
| | ACAAAGAATT | CTTTATTCAG | CATTCGACCT | AGTTGAACAA | CGCAGgtTCG | TGATGCATTA | 8820 |
| 35 | GAAGTATTCG | AAGAAGCAAT | CAACAACATT | ATGCCAGTAT | TAGAAGTTAA | AGCTCGTCGC | 8880 |
| | GTAGGTGGTT | CTAACTATCA | AGTACCAGTA | GAAGTTCGTC | CAGAGCGTCG | TACTACTTTA | 8940 |
| | GGTTTACGTT | GGTTAGTTAA | CTATGCACGT | CTTCGTGGTG | AAAAAACGAT | GGAAGATCGT | 9000 |
| 40 | TTAGCTAACG | AAATTTTAGA | TGCAGCAAAT | AATACAGGTG | GTGCCGTTAA | GAAACGTGAG | 9060 |
| | GACACTCACA | AAATGGCTGA | AGCAAACAAA | GCATTTGCTC | ACTACCGTTG | GTAAGATAAA | 9120 |
| | AGCTTTTACC | CTGAGTGTGT | TCTATATTAA | TGAATTTTCA | TTAAGCGTTC | ATGCTTAGGG | 9180 |
| 45 | CATCGCCATA | TCTATCGTAT | TTATTCAGTA | ATATAAACTG | GAAGGAGAAA | AAATACATGG | 9240 |
| | CTAGAGAATT | TTCATTAGAA | AAAACTCGTA | ATATCGGTAT | CATGGCTCAC | ATTGATGCTG | 9300 |
| | GTAAAACGAC | TACGACTGAA | CGTATTCTTT | ATTACACTGG | CCGTATCCAC | AArGknGGTG | 9360 |
| 50 | AAaCACACGA | AGGTGCTTCA | CAAATGGACT | GGATGGAGCA | AGAACAAGAC | CGTGGTATTA | 9420 |
| | CTATCACATC | TGCTGCAACA | ACAGCAGCTT | GGGAAGGTCA | CCGTGTAAAC | ATTATCGATA | 9480 |

| | CAGTTACAGT | ACTTGATGCA | CAATCAGGTG | TTGAACCTCA | AACTGAAACA | GTTTGGCGTC | 9600 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | AGGCTACAAC | TTATGGTGTT | CCACGTATCG | TATTTGTAAA | CAAAATGGAC | AAATTAGGTG | 9660 |
| 5 | CTAACTTCGA | ATACTCTGTA | AGTACATTAC | ATGATCGTTT | ACAAgCTAAC | GCTGCTCCAA | 9720 |
| | TCCAATTACC | AATTGGTGCG | GAAGACGAAT | TCGAAGCAAT | CATTGACTTA | GTTGAAATGA | 9780 |
| _ | AATGTTTCAA | ATATACAAAT | GATTTAGGTA | CTGAAATTGA | agaaattgaa | ATTCCTGAAG | 9840 |
| 10 | ACCACTTAGA | TAGAGCTGAA | GAAGCTCGTG | CTAGCTTAAT | CGAAGCAGTT | GCAGAAACTA | 9900 |
| ٠ | GCGACGAATT | AATGGAAAAA | TATCTTGGTG | ACGAAGAAAT | TTCAGTTTCT | GAATTAAAAG | 9960 |
| 15 | AAGCTATCCG | CCAAGCTaCt | ACTAACGTAG | AATTCTACCC | AGTACTTTGT | GGTACAGCTT | 10020 |
| | TCAAAAACAA | AGGTGTTCAA | TTAATGCTTG | ACGCTGTAAT | TGATTACTTA | CCTTCACCAC | 10080 |
| | TAGACGTTAA | ACCAATTATT | GGTCACCGTG | CTAGCAACCC | TGAAGAAGAA | GTAATCGCGA | 10140 |
| 20 | AAGCAGACGA | TTCAGCTGAA | TTCGCTGCAT | TAGCGTTCAA | AGTTATGACT | GACCCTTATG | 10200 |
| | TTGGTAAATT | AACATTCTTC | CGTGTGTATT | CAGGTACAAT | GACATCTGGT | TCATACGTTA | 10260 |
| | AGAACTCTAC | TAAAGGTAAA | CGTGAACGTG | TAGGTCGTTT | ATTACAAATG | CACGCTAACT | 10320 |
| 25 | CACGTCAAGA | AATCGATACT | GTATACTCTG | GAGATATCGC | TGCTGCGGTÅ | GGTCTTAAAG | 10380 |
| | ATACAGGTAC | TGGTGATACT | TTATGTGGTG | AGAAAAATGA | CATTATCTTG | GAATCAATGG | 10440 |
| | AATTCCCAGA | GCCAGTTATT | CACTTATCAG | TAGAGCCAAA | ATCTAAAGCT | GACCAAGATA | 10500 |
| 30 | AAATGACTCA | AGCTTTAGTT | AAATTACAAG | AAGAAGACCC | AACATTCCAT | GCACACACTG | 10560 |
| | ACGAAGAAAC | TGGACAAGTT | ATCATCGGTG | GTATGGGTGA | GCTTCACTTA | GACATCTTAG | 10620 |
| | TAGACCGTAT | GAAGAAAGAA | TTCAACGTTG | AATGTAACGT | AGGTGCTCCA | ATGGTTTCAT | 10680 |
| 35 | ATCGTGAAAC | ATTCAAATCA | TCTGCACAAG | TTCAAGGTAA | ATTCTCTCGT | CAATCTGGTG | 10740 |
| | GTCGŦGGTCA | ATACGGTGAT | GTTCACATTG | AATTCACACC | AAACGAAACA | GGCGCAGGTT | 10800 |
| 40 | TCGAATTCGA | AAACGCTATC | GTTGGTGGTG | TAGTTCCTCG | TGAATACATT | CCATCAGTAG | 10860 |
| | AAGCTGGTCT | TAAAGATGCT | ATGGAAAATG | GTGTTTTAGC | AGGTTATCCT | TTAATTGATG | 10920 |
| | TTAAAGCTAA | ATTATATGAT | GGTTCATACC | ATGATGTCGA | TTCATCTGAA | ATGGCCTTCA | 10980 |
| 45 | AAATTGCTGC | ATCATTAGCA | CTTAAAGAAG | CTGCTAAAAA | ATGTGATCCT | GTAATCTTAG | 11040 |
| | AACCAATGAT | GAAAGTAACT | ATTGAAATGC | CTGAAGAGTA | CATGGGTGAT | ATCATGGGTG | 11100 |
| | ACGTAACATC | TCGTCGTGGA | CGTGTTGATG | GTATGGAACC | TCGTGGTAAT | GCACAAGTTG | 11160 |
| 50 | TTAATGCTTA | TGTACCACTT | TCAGAAATGT | TCGGTTATGC | AACATCATTA | CGTTCAAACA | 11220 |
| | CTCAAGGTCG | CGGTACTTAC | ACTATGTACT | TCGATCACLA | TGCTGAAGTT | CCaAAATCaA | 11280 |

| | GCCTAGGTT | AAATACAAGO | TGAGCTTAA | TGTAAGCTAT | CATCTTTATA | GTTTGATTTT | 1140 |
|----|------------|-------------|------------|------------|------------|------------|-------|
| | TTGGGGTGA | A TGCATTATA | AAGAATTGTA | AAATTCTTTT | TGCATCGCTA | TAAATAATTT | 1146 |
| 5 | CTCATGATGO | TGAGAAACTA | TCATGAGAGA | AAATTTAAA | TATTATTTT | AATTAGAATA | 1152 |
| | GGAGAGATTT | TATAATGGCA | AAAGAAAAT | TCGATCGTTC | TAAAGAACAT | GCCAATATCG | 1158 |
| 10 | GTACTATCGG | TCACGTTGAC | CATGGTAAAA | CAACATTAAC | AGCAGCAATC | GCTACTGTAT | 11646 |
| 10 | TAGCAAAAA | TGGTGACTCA | GTTGCACAAT | CATATGACAT | GATTGACAAC | GCTCCAGAAG | 11700 |
| | AAAAAGAACG | TGGTATCACA | ATCAATACTT | CTCACATTGA | GTACCAAACT | GACAAACGTC | 11760 |
| 15 | ACTACGCTCA | CGTTGACTGC | CCAGGACACG | CTGACTACGT | TAAAAACATG | ATCACTGGTG | 11820 |
| | CTGCTCAAAT | GGACGGCGGT | ATCTTAGTAG | TATCTGCTGC | TGACGGTCCA | ATGCCACAAA | 11880 |
| | | | | | | GTATTCTTAA | 11940 |
| 20 | | | | | AGTAGAAATG | - | 12000 |
| | ACTTATTAAG | CGAATATGAC | TTCCCAGGTG | ACGATGTACC | TGTAATCGCT | GGTTCAGCAT | 12060 |
| | TAAAAGCTTT | AGAAGGCGAT | GCTCAATACG | AAGAAAAAT | CTTAGAATTA | ATGGAAGCTG | 12120 |
| 25 | TAGATACTTA | CATTCCAACT | CCAGAACGTG | ATTCTGACAA | ACCATTCATG | ATGCCAGTTG | 12180 |
| | AGGACGTATT | CTCAATCACT | GGTCGTGGTA | CTGTTGCTAC | AGGCCGTGTT | GAACGTGGTC | 12240 |
| | AAATCAAAGT | TGGTGAAGAA | GTTGAAATCA | TCGGTTTACA | TGACACATCT | AAAACAACTG | 12300 |
| 30 | TTACAGGTGT | TGAAATGTTC | CGTAAATTAT | TAGACTACGC | TGAAGCTGGT | GACAACATTG | 12360 |
| | GTGCATTATT | ACGTGGTGTT | GCTCGTGAAG | ACGTACAACG | TGGTCAAGTA | TTAGCTGCTC | 12420 |
| | CTGGTTCAAT | TACACCACAT | ACTGAATTCA | AAGCAGAAGT | ATACGTATTA | TCAAAAGACG | 12480 |
| 35 | AAGGTGGACG | TCACACTCCA | TTCTTCTCAA | ACTATCGTCC | ACAATTCTAT | TTCCGTACTA | 12540 |
| | CTGAÉGTAAC | TGGTGTTGTT | CACTTACCAG | AAGGTACTGA | AATGGTAATG | CCTGGTGATA | 12600 |
| 40 | | | | | | ACTCGTTTCT | 12660 |
| | CAATCCGTGA | AGGTGGACGT | ACTGTAGGAT | CAGGCGTTGT | TACTGAAATC | ATTAAATAAT | 12720 |
| | TTCTAATTTC | TTAGATTTTA | TATAAAAAGA | AGATCCCTCA | ATCGAGGGGt | CTTTTTTTAA | 12780 |
| 45 | TGTGTAAATT | TTGTAATGGC | TATTCGATTT | AGAAGAACAA | TAATTGATGA | AAGACTGACT | 12840 |
| | AATAAAACTT | ATAACTGATA | ATACTGTTTA | AATAAAATTG | TTGAGTCTTG | GACATTGTAA | 12900 |
| | AATGCTCCCT | | | | | | 12960 |
| 50 | TTTATGTCTC | | | | | | 13020 |
| | | | | | | TTTTCATCAA | 13080 |

| | TCAGAAAGAA | TGCACCTGGT | CGTACTITCA | AATAATGTGA | AAAATCTTCT | CCAATCATCA | 1320 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | TTAAATCTGA | TTCATTAAAG | CGTACATGTA | AGTCATTTGT | TGCTTCTTTA | ATAACTTGAT | 1326 |
| 5 | ATGCTTTCTC | GTTATTATGG | ACAGGCAAAT | ACCCTTTAAT | ATAATTCAAA | TCATAGTTAA | 13320 |
| | TATCATTTGC | TATTGCTAAA | CCTTGTAGAA | GCTTATCCAT | TTTGTCCATT | ACATGATTCT | 13380 |
| | GTATATCTGA | ATCGAAAGTT | CTAACTGTAC | CTTTACAAAA | TGCTTGATCA | GGAATAACGC | 13440 |
| 10 | TATCTGTGGT | GCCTGCTTGA | ATCATTCCAA | ATGAAAGTAC | AGCTTGTTTA | ACTGGATCGA | 13500 |
| | TCGTACGTGA | AATTATTTTT | TGTGCACTTA | AAATGAACTC | TGCCATGATT | ACTATTGGGT | 13560 |
| 15 | CAATGGTTTC | ATGAGGTTTG | GCACCATGAC | CACCACGACC | TTTAAATGTG | ACGCTAAATT | 13620 |
| | CATCTGGAGA | GGCCATGATT | GCCCCCGCAC | GTGAATGAAT | AGTTCCAGTA | GGATAACCAC | 13680 |
| | TCCATAAATG | TGTACCGTAA | ATTCTATCTA | CATTTTCCAG | ACATCCAGCA | TCTATCATTT | 13740 |
| 20 | CTTGAGAACC | ACCTGGCATG | ATTTCTTCAC | CGTACTGGAA | TATTAATACA | ACATTACCTT | 13800 |
| | CTAATAAATG | TTTATGTTCA | TCTAAAATCT | CTGCTACAGT | AAGTAAAATT | GCTGTATGAC | 13860 |
| | CATCATGCCC | ACACGCATGC | ATACATCCTG | GATTTTTAGA | CTTATAAGGC | ACATCGTTTA | 13920 |
| 25 | ATTCCTCGAC | AGGTAACGCA | TCAAAGTCAG | CTCTTAATGC | AATGGTAGGT | CCTGTGCCCA | 13980 |
| | AGCCTTTAAA | TGTGGCTTTG | ATACCATTGC | GGCCGATAGG | AGTTTCAATA | TCACAAGATA | 14040 |
| | ACTGGCTTAA | TTGGTTAACA | ATATAATCAT | GTGTTTGAAA | TTCTTCAAAA | GATAACTCAG | 14100 |
| 30 | GATATTGGTG | TAAATAACGT | CTGAGTTGAA | TTGTTTTATT | TTCTTTATTA | TTTGCTAGTT | 14160 |
| | GGAACCAATC | TAACACCCTT | ATCACTACTT | TCTAAAATAA | TGTTTATAGT | ATAACATTTT | 14220 |
| | ATGAAATTAT | CGTACTAAAT | GATTGCTTTG | AGATATTTTA | TCTATGAATG | ATAAGGCTTT | 14280 |
| 35 | CAAGTTATGT | AGAATTACTG | TATGATAAAG | GTATTACCAA | ACAATACTTA | AGGGGGATTA | 14340 |
| | TATACTGTGG | TTCAATCATT | ACATGAGTTT | TTAGAGGAAA | ATATAAATTA | TCTAAAAGAA | 14400 |
| 10 | AATGGTTTGT | ATAATGAAAT | AGATACAATT | GAAGGTGCAA | ACGGACCAGA | AATCAAAATC | 14460 |
| | AATGGGAAAT | CATACATTAA | CTTATCTTCA | AATAATTATT | TAGGACTAGC | AACAAATGAA | 14520 |
| | GATTTGAAAT | CaGctGCAAA | AGCAGCTATT | GATACACATG | GTGTAGGTGC | AGGCGCTGTT | 14580 |
| 15 | CGTACAATCA | ATGGTACATT | AGATTTACAC | GACGAATTAG | AAGAAACACT | AGCAAAATTT | 14640 |
| | AAAGGAACAG | AAGCTGCAAT | AGCTTATCAA | TCAGGATTTA | ATTGTAATAT | GGCTGCTATT | 14700 |
| | TCAGCTGTCA | TGAATAAAAA | TGATGCTATT | TTATCAGATG | AGCTTAATCA | TGCATCAATT | 14760 |
| 50 | ATTGATGGAT | GTCGCTTATC | TAAAGCTAAA | ATTATTCGAG | TTAACCATTC | AGACATGGAT | 14820 |
| | GATTTACGTG | CGAAAGCAAA | AGAAGCAGTT | GAATCAGGTC | AATACAATAA | AGTGATGTAT | 14880 |

| ATTGCAGAAG AATTTGG | TTT ATTAACTTAT | GTTGACGACG | CTCATGGTTC | AGGTGTTATG | 15000 | | | |
|------------------------------------|----------------|------------|------------|------------|-------|--|--|--|
| GGTAAAGGCG CTGGTAC | GGT TAAACATTTT | GGTTTACAAG | ATAAAATCGA | TTTCCAAATA | 15060 | | | |
| GGTACGCTTT CTAAAGC | AAT TGGTGTCGTT | GGCGGTTATG | TAGCAGGTAC | AAAAGAGTTA | 15120 | | | |
| ATAGATTGGT TAAAAGC | ACA ATCACGACCA | TTCTTATTCT | CTACATCATT | AGCACCTGGG | 15180 | | | |
| GATACCAAAG CAATAAC | TGA AGCAGTTAAA | AAGTTAATGG | ATTCAACTGA | ATTACATGAT | 15240 | | | |
| AAATTATGGA ACAATGC | ACA ATATTTAAAA | AATGGATTGT | CAAAATTAGG | ATATGATACA | 15300 | | | |
| GGTGAGTCAG AAACTCC | AAT TACACCAGTA | ATTATTGGTG | ATGAAAAAAC | AACTCAAGAA | 15360 | | | |
| TITAGTAAGC GTTTAAA | AGA CGAAGGTGTC | TATGTGAAAT | CTATCGTTTT | CCCAACAGTA | 15420 | | | |
| CCAAGAGGTA CAGGACG | TATAAADAA TDI | CCTACAGCTG | CACATACAAA | AGACATGTTA | 15480 | | | |
| GATGAAGCAA TTGCGGC | ITA TGAAAAGTA | GGAAAAGAAA | TGAAGTTGAT | TTAATATTTA | 15540 | | | |
| TTTATTCCCA CGGCAAA | TAT TGTCGTGGGC | TTTTTTTAAT | GTTTAGTTTA | TTAACAGT | 15598 | | | |
| (2) INFORMATION FOR SEQ ID NO: 83: | | | | | | | | |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 661 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 83:

| | AAGTAAATCA | ACTTACTGGG | ATAAGAATAA | AGGCGATTAT | AGTAACAAGT | TGATTTTATT | 60 |
|------|------------|------------|------------|------------|------------|------------|-----|
| | CGAAAAACAT | TTTGAACCGG | TTCTGGGTAT | CAAGATGCAA | CATAGTGGAG | GTCATAGCTT | 120 |
| 35 | TGGCCACACG | ATTATTACGA | TTGAAAGTCA | AGGAGATAAA | GCAGTTCATA | TGGGTGATAT | 180 |
| | ATTCCCAACT | ACTGCACATA | AAAATCCTCT | ATGGGTAACG | GCATATGATG | ATTATCCTAT | 240 |
| -40 | GCAATCGATT | CGTGAAAAAG | AACGCATGAT | ACCATATTTT | ATTCAGCAAC | AATATTGGTT | 300 |
| . ** | CTTGTTTTAT | CATGATGAAA | ACTACTTTGC | TGTAAAATAC | AGCGATAATG | GTGAAAACAT | 360 |
| | AGATGCATAT | ATTTTACGTG | AAACATTAGT | TGATAATAAC | TAAAATAAAG | ATGTATTACT | 420 |
| 45 | AAACAAATTT | TCAAAAATAA | AAAATTGAGC | CACATCCAAT | CTTACTAATT | AGGGTGTGGC | 480 |
| | TCATTTTTAA | GTTTTACGAT | CCAAATCAAA | TATGGaTAAA | ATTCgTATTA | ACGCTCTACa | 540 |
| | ATGETAATGA | CTTCACCAGT | ATATGCATCT | GCATAAAAAT | CATAATGAAT | ATTTTGACCA | 600 |
| 50 | TTTTTAATAG | TTGTAATTCC | ACCTTGATAA | ACTAAACGGT | ATTTATCAGT | TTCAGGATGA | 660 |
| | A | | | | | | 661 |

55

5

10

15

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5738 base pairs (B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 84:

| 60 | GGTGCATTCT | GAATCTAATC | AaACTGTAGT | AGTCGCACCA | CAGCAGTTAA | GCAGACGGTA | • |
|-------|------------|------------|-------------|------------|------------|------------|---|
| 120 | GATAAGCGAA | AGTAATTTTC | TCTTCTTCAA | CTTATATATA | AGTTGTCGCG | TTTTAGGATT | |
| 180 | TCAATTCAAA | TGTATTGGGT | TAGGATTGCC | GAGAAAGAAT | AGAAGATGTA | TTAAAGATGA | |
| 240 | ACAACACTAT | AAATACGACA | CAAAAAAGGA | CTACTTATGT | AGGATGGTTG | AATTTAATTA | |
| 300 | CGTTCAAACA | TCGAGGTATA | GTGAAAAGTT | TCAACAATTA | AAAACCAAAA | TTGTATATGA | |
| 360 | GAAAAGCCTG | GGTTACTTCT | AGCGCTTATT | GGTGAAGTAA | AAAAGCAAAT | TCATGTTTTC | |
| 420 | GCAGGCTATA | TTATGCACAA | TAGCGATTAC | GTATCGAATG | AAGTACAGTT | GTGCAGGTAA | |
| 480 | TTTAATGAGC | AAACTATATT | AgcCAACACA | GATATGCGTA | TATTGATGGC | AGACATTAGT | |
| 540 | GAAGCAATTA | GACTATGTCA | TTGGTCGAAC | AGCTTAATCA | TGGACTATCA | AAAATAATAA | |
| 600 | CCAAATCCAT | CCCTGTACCT | TAACAGCTGG | TTAGATTTGC | AATTGAAAAT | CGTCGACAGA | |
| 660 | AAACGTTACG | TCTGTTTAAT | AATTAGTTGA | AGGTTCAAAG | TGGGTCTGAA | CTGAGTTAAT | |
| . 720 | CTATATGCGC | TGATGCACAA | ATACTGTGAC. | CCGCCAGTTA | TGTCGATACA | ACATTATTAT | |
| 780 | AATGAAGTTA | AAATGATAAr | ATAGTGAAAA | TTAGTAATTG | AGATAGTCTG | GTGCTATTAA | |
| 840 | ATTTTGAACA | TCTAGGTGTC | GCAGTAACAT | GAAAAAGCAG | AGCACTTATG | AAAAAGCAAA | |
| 900 | TAAGTATGAT | TGGAGATGAA | ATCACTATTA | TCTAGTTATT | CGATAAATCT | AGACAAAGGT | |
| 960 | AAACAGAGAT | CCGACAAATG | CGATGACGGT | TGCCTAATAT | AACCATATAT | TGATĀTTCAT | |
| 1020 | TAACATCACA | GAAATCATTG | AGGTGTTACA | CGACAACACA | TTAAAACAAG | GATGGATCTT | |
| 1080 | TAAACCATAT | AAATCATGTT | AGAAAAAGTG | CCACACCTAT | CCTCGATATA | TCACTTACAT | |
| 1140 | AGGAAATAAG | TATTATGGTC | TCTAAAGTTT | AAGCACTAAA | GAGGAAGTAC | TGAAAGCTTA | |
| 1200 | TTAATGATTC | ATTAACGGTA | TCGAAAAGTT | ATGATATTGA | CAAATCCTTA | AATTACCGAT | |
| 1260 | ATCAATTATE | CACTATACTG | TGAAGTTCCA | TTCCATCAAA | CTAATAGAAT | ACGCTATTTA | |
| 1320 | GGAATAAAGC | CATCCAGAGC | GATTATTGCA | GCTTTGTACC | CAGAGTAAAG | TTTCGAATtA | |
| 1380 | TAAGTCAAGT | AAAGGTGCTT | TTTAATTAAC | TACTATACGA | AACCTTGACA | AATAAGTCAA | |
| 1440 | CAATTCAAAT | AGAAAATTAG | TAAAAAAATT | GTATTTCCGG | TCATTAGCGG | GACAACGGcG | |
| | | | | | | | |

| | GTTCTTAATG | AAAGACTTAT | TTAATGATAA | GAAATTACGT | GATTATTATG | AAGATATGAA | 1560 |
|----|------------|--------------|------------|------------|------------|------------|------|
| | CGGATTTATT | AGTAATGCGA | AGTTAGTTGT | TGATGATAAA | AAAATTCCTA | AACGAATGCC | 1620 |
| 5 | ACAACAAGAT | TATAAACAGA | AAAGATGGTT | TGGGTTATAA | ACAGCAAATG | AGGGGTTTTA | 1680 |
| | TGGCACATTT | ATCTGTGAAA | TTGCGGCTTT | TAATACTAGC | ATTAATCGAT | TCACTGATAG | 1740 |
| | TGACATTTTC | AGTATTCGTA | AGTTATTACA | TTTTAGAACC | GTATTTCAAA | ACATATTCTG | 1800 |
| 10 | TCAAATTATT | ' AATATTGGCA | GCTATATCAC | TATTCATATC | GCATCATATT | TCaGCATTTA | 1860 |
| | TTTTTAATAT | GTATCATCGA | GCGTGGGAAT | ATGCCAGTGT | GAGTGAATTG | ATTTTAATTG | 1920 |
| 15 | TTAAAGCTGT | GACGACATCT | ATCGTTATTA | CGATGGTGGT | CGTGACAATT | GTTACAGGCA | 1980 |
| | ATAGACCGTT | TTTTAGATTG | TATTTAATTA | CTTGGATGAT | GCACTTGATT | TTAATAGGTG | 2040 |
| | GCTCAAGGTT | ATTTTGGCGT | ATTTATCGGA | AATACCTTGG | AGGTAAGTCA | TTTAATAAGA | 2100 |
| 20 | AGCCAACTTT | AGTTGTTGGT | GCTGGTCAAG | CAGGTTCAAT | GCTGATTAGA | CAAATGTTGA | 2160 |
| | AAAGTGACGA | AATGAAACTT | GAACCGGTAT | TAGCAGTCGA | TGATGACGAA | CATAAACGCA | 2220 |
| | ATATCACAAT | TACTGAGGGT | GTAAAAGTCC | AAGGTAAAAT | TGCGGATATT | CCAGAACTAG | 2280 |
| 25 | TGAGGAAATA | TAAGATTAAA | AAAATCATCA | TTGCAATTCC | AACTATTGGT | CAAGAGCGTT | 2340 |
| | TGAAAGAAAT | TAATAATATT | TGCCATATGG | ATGGCGTTGA | GTTATTGAAA | ATGCCAAATA | 2400 |
| | TAGAAGACGT | CATGTCTGGT | GAGTTAGAAG | TGAACCAACT | TAAAAAAGTT | GAAGTAGAAG | 2460 |
| 30 | ATTTACTAGG | CAGAGATCCT | GTTGAATTAG | ATATGGATAT | GATATCAAAT | GAATTGACGA | 2520 |
| | ATAAAACTAT | TTTAGTTACG | GGTGCAGGTG | GTTCAATAGG | ATCAGAAATT | TGTAGACAAG | 2580 |
| 35 | TTTGTAATTT | CTATCCAGAA | CGTATTATTC | TACTTGGCCA | TGGTGAAAAC | AGTATTTATT | 2640 |
| | _ | | | | TGATATCGTT | | 2700 |
| | CGGATGTGCA | AAATAGAGCG | CGTATGTTTG | AAATTATGGA | AACGTATAAA | CCATACGCAG | 2760 |
| 10 | TTTATCATGC | AGCAGCACAC | AAGCACGTGC | CGTTAATGGA | AGACAACCCT | GAAGAAGCAG | 2820 |
| | TACGTAATAA | TATTTTAGGT | ACGAAAAATA | CTGCTGAAGC | TGCTAAAAAT | GCAGAGGTAA | 2880 |
| | | | | | | ATGGGAGCTT | 2940 |
| 15 | CAAAGCGAAT | TGCAGAAATG | ATTATTCAAA | GTTTAAATGA | TGAAACGCAT | CGAACAAATT | 3000 |
| | TTGTTGCAGT | GAGATTTGGT | AATGTACTTG | GATCGAGAGG | ATCTGTGATT | CCACTTTTCA | 3060 |
| | | | | | | ACACGTTACT | 3120 |
| 60 | TTATGACAAT | TCCTGAAGCT | TCTAGACTAG | TTTTGCAGGC | AGGGGCATTA | GCAGAAGGTG | 3180 |
| | CCCSSCOSSO | mamaamsass | | | | | |

| | CCGGCGAAAA | AATGTTTGAA | GAGCTTATGA | ATAAAGATGA | GGTTCATCCT | GAACAAGTAT | 3360 |
|-----|------------|------------|------------|------------|------------|-----------------|------|
| _ | TTGAAAAAAT | TTATCGTGGC | AAAGTACAAC | ATATGAAATG | TAATGAAGTT | GAAGCGATTA | 3420 |
| 5 | TTCAAGACAT | CGTCAATGAC | TTTAGTAAAG | AAAAAATTAT | TAACTATGCC | AATGGCAAAA | 3480 |
| | AGGGAGATAA | TTATGTTCGA | TGACAAAATT | TTATTAATTA | CTGGGGGCAC | AGGATCATTC | 3540 |
| 10 | GGTAATGCTG | TTATGAAACA | GTTTTTAGAT | TCTAATATTA | AAGAAATTCG | TATTTTTCA | 3600 |
| , 0 | CGCGATGAGA | AAAAACAAGA | TGACATTCGA | AAAAAATATA | ATAATTCAAA | ATTAAAGTTC | 3660 |
| | TACATTGGTG | ATGTGCGTGA | TAGTCAAAGT | GTAGAAACAG | CAATGCGAGA | TGTTGATTAC | 3720 |
| 15 | GTATTCCATG | CAGCAGCTTT | AAAACAAGTG | CCGTCATGTG | AATTCTTTCC | AGTTGAGGCA | 3780 |
| | GTGAAGACAA | ATATTATTGG | TACAGAAAAT | GTCTTACAAA | GTGCTATTCA | TCAAAATGTT | 3840 |
| | AAAAAAGTCA | TATGTTTATC | TACAGATAAG | GCAGCGTATC | CTATTAATGC | TAGGGGTATT | 3900 |
| 20 | TCAAAAGCAA | TGATGGAAAA | AGTATTCGTA | GCCAAATCAA | GAAATATTCG | TAGTGAACAA | 3960 |
| | ACGCTTATTT | GTGGTACAAG | ATACGGTAAT | GTGATGGCTT | CAAGAGGATC | AGTAATACCT | 4020 |
| | TTGTTTATCG | ACAAAATCAA | AGCTGGAGAA | CCTTTAACGA | TTACAGATCC | TGATATGACA | 4080 |
| ?5 | AGATTTTTAA | TGAGCTTAGA | AGATGCGGTA | GAACTAGTTG | TTCATGCATT | TAAGCATGCA | 4140 |
| | GAGACAGGAG | ATATTATGGT | TCAAAAAGCA | CCAAGCTCAA | CGGTAGGGGA | TCTTGCGACC | 4200 |
| | GCATTATTAG | AATTGTTTGA | AGCTGATAAT | GCAATTGAAA | TCATTGGTAC | GCGACATGGA | 4260 |
| 30 | GAGAAAAAG | CAGAAACATT | GTTGACGAGA | GAAGAATACG | CACAATGTGA | AGATATGGGT | 4320 |
| | GATTATTTTA | GAGTGCCGGC | AGACTCCAGA | GATTTAAATT | ATAGTAATTA | TGTTGAAACC | 4380 |
| 35 | GGTAACGAAA | AGATTACGCA | ATCTTATGAA | TATAACTCCG | ATAATACACA | TATTTTAACG | 4440 |
| | GTGGAAGAGA | TAAAAGAAAA | ACTTTTAACA | CTAGAATATG | TTAGAAACGA | ATTGAATGAT | 4500 |
| | TATAAAGCTT | CAATGAGATA | GGAGAGATTG | ACGTTGAATA | TTGTAATTAC | AGGAGCAAAA | 4560 |
| 10 | GGTTTTGTAG | GAAAAAACTT | GAAAGCAGAT | TTAACTTCAA | CGACAGATCA | TCATATTTTC | 4620 |
| | GAAGTACATC | GACAAACTAA | AGAGGAAGAA | TTAGAGTCAG | CATTGTTGAA | AGCAGACTTT | 4680 |
| | GTCGTGCATT | TAGCGGGTGT | TAATCGACCT | GAACATGACA | AAGAATTCAG | CTTAGGAAAC | 4740 |
| 15 | GTGAGTTATT | TAGATCATGT | ACTTGATATA | TTAACTAGAA | ATACGAAAAA | GCCAGCGATA | 4800 |
| | TTATTATCGT | CTTCAATACA | AGCAACACAA | GATAATCCTT | ATGGTGAGAG | TAAGTTGCAA | 4860 |
| | GGGGAACAGC | TATTAAGAGA | GTATGCCGAA | GAGTATGGCA | ATACGGTTTA | TATTTATCGC | 4920 |
| 50 | TGGCCAAATT | TATTCGGCAA | GTGGTGTAAG | CCGAATTATA | ACTCAGTGAT | AGCAACATTT | 4980 |
| | TGTTACAAAA | TTGCACGTAA | CGAAGAGATT | СВАСТТВАТС | ATCCCAATCT | TC A A CTA A CC | E040 |

| | | • | | | ** | |
|------------|---|--|--------------|------------|------------|-----------|
| ATTGAAAATG | GTGTACCTAC | AGTACCAAAC | GTATTTAAAG | TGACATTGGG | AGAAATTGTA | - 5160 |
| GATTTATTAT | ACAAGTTCAA | ACAGTCACGT | CTCGATCGAA | CATTGCCGAA | ATTAGATAAC | 5220 |
| TTGTTTGAAA | AAGATTTGTA | TAGTACGTAT | TTAAGCTATC | TACCTAGTAC | aGACTTTAGT | 5280 |
| TAYCCCTTAC | TTATGAATGT | GGATGATAGG | GGTTCTTTTA | CAGAATTTAT | AAAAACACCG | 5340 |
| GATCGTGGTC | AAGTTTCTGT | AAATATTTCT | AAACCAGGTA | TTACTAAAGG | TAATCACTGG | 5400 |
| CATCATACTA | AAAACGAAAA | ATTTCTAGTC | GTATCAGGTA | AAGGGGTAAT | TCGTTTTAGA | 5460 |
| CATGTTAATG | ATGATGAAAT | CATTGAATAT | TATGTTTCTG | GCGACAAATT | AGAAGTTGTA | 5520 |
| GACATACCAG | TAGGATACAC | ACATAATATT | GAAAATTTAG | GCGACACAGA | TATGGTAACT | 5580 |
| ATTATGTGGG | TGAATGAAAT | GTTTGATCCA | AATCAGCCAG | ATACGTATTT | CTTGGAGGTA | 5640 |
| TAGCGCATGG | aaaaactgaa | rTTAATGACA | ATAGTTGGTA | CAAGGCCTGA | AATCATTCGT | 5700 |
| TTATCATCAA | CGATTAAAGC | ATGTGATCAA | TATETTAA | | | 5738 |
| (2) INFORM | ATION FOR SE | Q ID NO: 85 | 5: | | | |
| | EQUENCE CHAR (A) LENGTH: (B) TYPE: nu (C) STRANDED (D) TOPOLOGY | 9062 base p cleic acid NESS: doubl | pairs | | | |
| (xi) 5 | SEQUENCE DES | CRIPTION: S | SEQ ID NO: 8 | 35 : | | |

| AICHICHACH A | CANTONIAL | IIIICCCAIC | IACIAIAICI | TITACCGCAG | ATAACTICAC | 60 |
|--------------|------------|------------|------------|------------|------------|-------|
| TCTCACACCT T | GCTCACGTA | ATTCTTGAGT | TGGTTGAATA | AATGTTCTTG | CAACATATTG | 120 |
| ATTTTTAACT A | GTCCCATTT | CATATGGCAA | ACCTATTTCT | TCAGCATAAC | CACTCGCAGC | 180 |
| TGAŢĀGCGAT g | AATTGGGTA | CACCGATGAC | CATATCAGCA | TTTACAGGGC | TTTCTTGGGC | 240 |
| TAATTTTTTA_C | CAGAAGCTT_ | TACGTACTGC | ATGGACATTT | TTACCAGCTA | TTGTTGAGTC | 3.0.0 |
| TGGTCTAGCA A | AATAAATAT | ATTCCATCGC | AGAAATTGCA | GTTGTCGTAT | GATGTGTATA | 360 |
| AGATTTAACT G | TAATACCTT | TATCGTTAAT | CACGACATAT | TCACCTGCAT | GAATATCTTG | 420 |
| AACAAATTCT G | CACCTAACA | CATCTATTGC | ACATGTTTCA | CTTGCAAGGA | TGTATGTCCC | 480 |
| ATCTTTCATT T | TACCTACAA | CAAGTGGTCT | GATAGCATTT | GGATCTACTG | CGCCATATAA | 540 |
| CGCATCTTTA G | TTAAAATCG | CAAATGTAAA | ACCGCCTTTA | ACTTTTCGCA | AACTITCTTT | 600 |
| CAACGCTTCC T | CAAAAGTAG | GAGCTTTACT | TCGACGTATC | AAATGCATAA | TGACTTCAGT | 660 |
| ATCAGAAGAC G | AATGGAAGA | TAGCACCTTG | TTTTTCTAAA | TTCTGACGCA | ATGATTTAGC | 720 |

| | CGGTTGAATA | TTTTCAATAC | CTTTATTACC | TGAAGTAGCA | TAACGGACGT | GACCAATTGC | 840 |
|----|--------------------|------------|------------|-------------|----------------|-------------|------|
| | ATGTTGATAT | CCTTTTAATC | GTTCCATTTG | ATCATCTTTA | ATCGCTTCAG | TTAGTAAGCC | 900 |
| 5 | TAATCCTCGC | TCGCCTTTTA | ATTCATTTTG | ATCAGAAACA | ACTATACCTG | CACCTTCTTG | 960 |
| | ACCACGATGT | TGCAAACTAT | GAAGTCCCAT | ATACGTTAGT | TGCGCTGCtT | CaGGATGATT | 1020 |
| | CCAAATACCA | AACACGCCAC | ATTCTTCGTT | TAATCCTGAG | TAGTTAAACA | TTGaGCAATT | 1080 |
| 10 | GCCCCtTCCC | ATATTTGTTT | AATATCTGAA | ACATTTTCAC | TAATCTCTGT | aTATGGTGTT | 1140 |
| | GTTACCTTGr | aATTATCACT | ATCTGTTAAA | AGTCCAATTT | CTATTGCATT | ATCAATATTT | 1200 |
| 15 | AAAGTTTTAC | CTGATTTAAC | AGAAACAACA | TATCGGCCTT | GCGTCTCACT | AAACAATTGT | 1260 |
| | GCATTTGTTA | TATCTATTGA | AGATTTTAAT | CCTAAACCGT | AATGCGCACT | TAGTTTAGCT | 1320 |
| | AAGGTAATCA | GTAAGCCACC | TTTACCAACT | GTTTGAACAT | GTGATAATAG | TCCTTCACGA | 1380 |
| 20 | ATAGCGGTCT | TGATTGATTC | ACCTTTTTCA | ACTTCTGAAC | TCAAATCTAA | TGACTCAAAT | 1440 |
| | TCATGATTAA | CTTTGCCATA | AATTAACTTT | TCAAGTTGAC | TACCACCAAA | GTCGTCCTTA | 1500 |
| | GTATCACCGA | TTAAATATAA | TTTATCTCCA | ACTTGAGGTT | CAAAATCATT | TAAATAATTT | 1560 |
| ?5 | ACATITICAA | TCAAACCTAC | CATTCCAACA | ACTGGTGTTG | GGAAAATAGA | AGTACCTTTC | 1620 |
| | GTTTCGTTAT | ATAAAGATAC | ATTACCAGAA | ACTACTGGTG | TCTTAAGAAT | GTCGCATGCT | 1680 |
| | TCTGCCATAC | CTTTCGTTGA | ATCTATCAAC | TGTTGATAĞA | TTTCTTTCTT | TTCAGGAGAA | 1740 |
| 30 | CCATAATTTA | AACAATCTGT | CATTGCTAAT | GGTGTTGCAC | CCACGGCAAT | TAAATTTCGA | 1800 |
| | TAAGCTTCAG | CTACTACCAT | CTTTCCACCT | TCATATGGAT | TGTTATATAC | ATAACGCGCT | 1860 |
| | TCACCATCAA | TTGTTGAAGC | AATTGCCTTA | TTTGTGCCTT | CCACACGTAC | TACCGATGCT | 1920 |
| 35 | TGAAGTCCTG | GCTTAATTAT | CGTATTGGCA | CCAACTTGTT | GGTCGTATTG | ATCATATAAA | 1980 |
| | TAGTGTTTAG | ATGCTATAGT | CGGATGCTTA | AGTAATTTAA | AGAAAGTATC | TTTAACATCG | 2040 |
| 10 | ATGTGTGTAT | AATCATTTTT | AGAAGTATTA | TAATCTTTTT | CTTCTCCTTC | TAAAATATAT | 2100 |
| | ACAGGTGCTT | CATCAGCTAG | TGGTTCAACT | GGAATGTCAG | CATAAACTTC | GTCATCATAT | 2160 |
| | GTTAAAACAA | AACGATTTGT | ATCTGTAACT | TCACCTATAA | CAGCACTATC | CAATTCGTGC | 2220 |
| 15 | TTATCAAATA | AATCTAAGAA | TTTTTGTTCA | GTACCTTTTT | CAACAACTAG | TAACATACGT | 2280 |
| | TCTTGAGTTT | CTGAAAGCAT | CATTTCATAA | GGAGAAATAC | CTGGCTCACG | TGTTGGCACT | 2340 |
| | TGTTCTAATC | TCAAATGTAA | CCCACTACCA | CCTTTTGCCG | CCATTTCAGA | CGATGAAGAT | 2400 |
| 50 | GTTAAACCAG | CAGCACCCAT | ATCTTGAATA | CCAACTAATT | CATCAAATGT | AATTGCTTCA | 2460 |
| | ACTION COMP | ССВТТВВТТТ | TTTACCTACA | NATION TO A | CC & Terrory C | ACA ACOTOOT | 2520 |

| | | - |
|------|---|------|
| | CGACCAGTTT TCAAACCAAC ATAAATGACC GAATTACCTA CACCTTTTGC TGTGCCTTTT | 2640 |
| 5 | TGAATCATGT CGTGATTGAT AACACCAACA CACATTGCAT TAACAAGTGG ATTGCCATCA | 2700 |
| J | TAACGTTCAT CAAATTCGAT TTCACCAGCA GTTGTTGGAA TACCAATGCA GTTACCATAA | 2760 |
| | CCTCCGATAC CCTTTACAAC ACCTTTAAGT AATCTTTGGT TTTGTTTATT ATCTAATTCT | 2820 |
| . 10 | CCAAATCTAA GACTGTTTAA CAAATTAATA GGTCTAGCCC CAATAGAGAC AATGTCACGA | 2880 |
| | ATGATTCCAC CAACGCCTGT AGCAGCCCCT TGATATGGTT CAATTGCTGA TGGATGATTG | 2940 |
| | TGAGACTCTA CTTTAAATAC TACGGCTTGA TTATCACCTA TATCGACTAC CCCTGCACCT | 3000 |
| 15 | TCACCAGGCC CCATAAGCAC ATGGTCACCT GACGTAGGAA ATTGCTTTAA AAACGGTTTA | 3060 |
| | GAATGTTTAT AAGAGCAATG TTCACTCCAC ATAACAGAAA AGATACCTGT TTCTGTAAAG | 3120 |
| | TTAGGTTGTC TGCCTAAAAT ATCGCAAACT TTTTCATATT CTTGATCaCT TAATCCCATA | 3180 |
| 20 | TCTTGATATA CTTTTTCAAG TTTAATTTCT TCAACGCTTG GTTCGATAAA TTTAGACATG | 3240 |
| | TTGTTCCCTC CAACTTTTTA CCATCGCTTC AAATAATTTC ACACCACTAT CAGTACCTAA | 3300 |
| | CAACGTTTCT AAAGCTCTTT CagGATGtGG CATCATGCCA CATACATTGC CTTTTTCGTT | 3360 |
| 25 | AACAATTCCT GCAATATCAT CATATGAACC GTTCGGATTA TTCACATATT TCAGAATAAT | 3420 |
| | TTGATTGTTA GCTTTTAATT GTTGATATAT TTCATCAGTA CAATAATAAT GACCTTCACC | 3480 |
| | GTGAGCTACA GGATATATAA CTTTTTCACC TTGTTCATAA AGATTTGTAA ATGCCGTTTG | 3540 |
| 30 | ATTATTCACT ATTTCTAACT CTTCATTTCT ACTAATAAAT AAATGTGAAT CGTTATGCAA | 3600 |
| | TAATGCACCA GGTAATAAGC CTATTTCAGT TAAAATTTGA AACCCATTAC AAACACCTAA | 3660 |
| | TACTGGCTTA CCTTCAGCTG CAAGACGTTT AACTTCCGAA ATAATCGGSG CTACACTAGC | 3720 |
| 35 | CATTGCCCCA GATCTTAAGT AATCCCCGAA TGAAAATCCA CCAGGAATAA GTACGCCATC | 3780 |
| | AAATÉCACTT AGTGATGTTT CTCTATAATC TACATATTCC GCTTCAACAC CACTTTTAAT | 3840 |
| -40 | AGCAGCATTA AACATGTCTC TATCACAATT CGAACCTGGA AAAACAAGAA CCGCAAATTT | 3900 |
| 40 | CATTITATGC ATTCTCCTTT TCATCATCTA ACACTITATA GCTATATTCT TCAATCACTG | 3960 |
| | TATTIGCAAA CAATTITICA CTIAGAGIIG TAATAAIGII GIGTACCIII TCATCACTAA | 4020 |
| 45 | CCTCATCCAC TGTCATATAT AATACTTTTC CTACACGAAT ATCATTCACT TGTGCATAAC | 4080 |
| | CTAAGTCATG TACAGCTCGA GTAAGCGTTT GTCCTTGCGT ATCTAATACT TGTGGTTGTA | 4140 |
| | ATGTGATATG TAGTTCAATT GTTTTCATTA TTTTAAATCC TCCAATTTGT TTAAAAATAT | 4200 |
| 50 | TTGATATGTT TCAATCAGTG ATCCAGTGTT ATTTCTATAT ACATCTTTAT CAAAGTTTGC | 4260 |
| | ATTGGTAGCT TTATCCCAAA TTCGACATGT ATCTGGAGAT ATTTCATCCG CTAACAAAAT | 4320 |

| | ATCCATTAAT | TGTTTCAACA | CATTATTAAT | CTTTAATGCT | TTGGATTTTA | GTATTTCAAT | 4440 |
|----|-------------|------------|------------|------------|------------|------------|------|
| | ATCTTCATCT | GATGCTATAT | TGAGCAATTT | AACATGGTCA | TCCGTTATCA | ACGGATCATT | 4500 |
| 5 | TAACGCATCA | TTTTTATAGA | AAAATTCTAC | AAGTGGTTCT | CTAAAAACTT | CACCATTTTC | 4560 |
| | AAAACCTAAA | CGCTTTGTAA | TAGATCCACT | AGCAATATTA | CGAACAACTA | CTTCTAATGG | 4620 |
| | AATTATTTTC | ACAGGCTTAA | CTAATTGTTC | TGTTTCAGAT | AATTGTTTAA | TAAAGTGACT | 4680 |
| 10 | TTCTATTCCA | TTTTCTTGTA | AATATTTAAA | TATAATAGAA | GTAATTTGAT | TATTTAATCG | 4740 |
| | CCCCTTACCT | GCCATTGTGT | CTTTCTTAGC | CCCGTTTCCA | GCAGTAACTT | CATCTTTATA | 4800 |
| 15 | TTCAACTCTT | AATTCATTTT | CTTGATTTGT | TGAGAAAATG | CGcTTCGCTT | TTCCTTCATA | 4860 |
| | TAATAATGTC | ATGCTTTAAT | TACTCCCCTC | AAATTTAGCG | TACATATCTT | GTTCAGTTTG | 4920 |
| | GTTTACATCA | TTCGTTAGTA | CAGTCATATG | CCCCATTTTT | CTGCTATCTT | TACGCTCAGA | 4980 |
| 20 | CTTACCATAA | ATATGTAAGT | GCCACTCTGG | ATGTTCATTA | AATTCATTTT | CCAATAAATC | 5040 |
| | TAAATCTTTA | CCTAGTAAGT | TCATCATGAC | TGCTGGCTTT | AATAATŢCAA | TTGAATTTGG | 5100 |
| | TAATGATTGT | CCGGTAACTG | CTAAAATATG | AGTATCAAAT | TGTGAATAAT | CACATGCTTC | 5160 |
| 25 | AATTGAATAA | TGTCCGGAAT | TGTGAGGCCT | TGGTGCTATC | TCGTTCACAT | ACAATTGGTT | 5220 |
| | GTTACTATCT | ATAAAAAATT | CAACTGTAAA | TGTTCCAATG | AAATGAATCG | ATTGGATAAT | 5280 |
| | TTTATTAACT | TGCTCTTTCG | CCTCAGCTGT | TTTATCTATT | CTCGCTGGAA | CAATTGTTTT | 5340 |
| 30 | GAAAAGTATT | TGATTTCTAT | GCTCATTTTC | TTGTAATGGG | AAAAAGTGA | TTTGATTGTT | 5400 |
| | GTTTCCTCTT | GTAACAGTAA | GAGATACTTC | TTTCTTGATA | TTCAAATATT | TTTCAGCTAC | 5460 |
| | GCATTCACTA | GTTTCAATTA | ATTTAAAACC | TTCTTGTAAG | TCTTTTTCGT | TGTTAATTAA | 5520 |
| 35 | AACTTGACCT | TTGCCATCGT | AGCCACCAAA | TCTAGTTTTT | ACAATAAAAG | GATATCCTAA | 5580 |
| | TGTTTCAATT | GCTTTGTCAA | TATCTGTAGA | TTCTTTTACT | GAAATGAACG | GGACAACTTT | 5640 |
| 40 | GGTACCAGCA | CTTTTTAATG | TTTCTTTTTC | AGTTAAGCGA | TCTTGTAATA | ACTGTATAGC | 5700 |
| | TIGGTAACCT | TGCGGAATAT | TGTACTTTTC | ACATAATAGT | TTTAATTGTT | GGGCTGAAAT | 5760 |
| | GTTTTCAAAT. | TCATAAGTAA | TCACATCACA | TTTTTGTCCT | AATTGATTGA | GTGCCTTTTC | 5820 |
| 45 | ATCGTCATAC | TTGGCTTGTA | TAAATTCGTG | TGCAACGTAT | CTACATGGAC | AATCTTCAGA | 5880 |
| | AGGATCCAAT | ACAACCACTT | TATAACCCAT | TTTTTGAGCT | GATTGTGCCA | TCATCTTTCC | 5940 |
| | AAGCTGACCA | CCACCAATAA | TGCCAATAGT | CGCACCAAAC | TTTAATTTAT | TGAAGTTCAT | 6000 |
| 50 | TTTGCATGTC | CTCCACTTTT | TGAATTAACG | AAGATTCATA | CTGATTTAGT | TTTTCAACTA | 6060 |
| | AAGAAGGATT | TTGAATACTT | AACATTCTTG | CTGCAAGTAT | ACCTGCGTTT | TTAGCACCTG | 6120 |

| | AAGAATCTAT ACCCTTTAAA CTTTTTGTTT CAATCGGCAC TCCAATAACT GGTAGCGTCG | 6240 |
|----|---|------|
| | TTAATGATGC AACCATACCT GGTAAATGTG CCGCACCGCC AGCGCCTGCA ATGATAATGT | 6300 |
| 5 | TTATACCTCT TTCTCTCGCT TCAGAAGCAA ATTGAACCAT CATTTTTGGC GTACGATGTG | 6360 |
| • | CGGATACTAC TIGTTTTTCG TACGGAATTT CAAAATAATC CAACATGTTA CAACTCTCTT | 6420 |
| 10 | GCATAATTTT CCAATCGGAA GAACTGCCCA TAATGACTGC TACTTTCACT TTGTACACCC | 6480 |
| ,, | TTTCAAAAGT TTGAATTGTG AATTACTTTA GTTGTATATT ATAGATATAG CATAACAAGC | 6540 |
| | AATTTCTGCT TTTTCAATCA AAAATCGAAC TTTATTTTGA TTTTTTATTT GAATTTACGT | 6600 |
| 15 | CTTTTGCTAT GTAAATTAGT TTTATAAACT AACAAAGTTA GGATATTGAC AATAGGAGGA | 6660 |
| | GAAGTITTTA TGGTTGCTAA AATTITAGAT GGTAAACAAA TTGCCAAAGA CTACAGACAG | 6720 |
| | GGGTTACAAG ATCAAGTTGA AGCGCTAAAA GAAAAGGGTT TTACACCTAA ATTATCCGTT | 6780 |
| 20 | ATATTAGTTG GTAATGATGG CGCTAGTCAA AGTTATGTTA GATCAAAAAA GAAAGCAGCT | 6840 |
| | GAAAAATTG GTATGATTTC AGAAATCGTA CATTTGGAAG AAACAGCTAC TGAAGAAGAA | 6900 |
| | GTATTAAACG AACTAAATAG ACTAAATAAT GATGATTCTG TAAGTGGTAT TTTGGTACAA | 6960 |
| 25 | GTACCATTAC CAAAACAAGT TAGCGAACAG AAAATATTAG AAGCAATCAA TCCTGAAAAA | 7020 |
| | GATGTGGACG GTTTTCATCC AATAAATATA GGGAAATTAT ATATCGATGA ACAAACTTTT | 7080 |
| | GTACCTTGCA CACCGCTCGG CATCATGGAA ATATTAAAAC ATGCTGATAT TGATTTAGAA | 7140 |
| 30 | GGTAAAAATG CAGTTGTAAT TGGACGAAGT CATATTGTCG GACAACCAGT TTCTAAGTTA | 7200 |
| | CTACTTCAAA AAAATGCATC AGTAACAATC TTACATTCTC GTTCAAAAGA TATGGCATCA | 7260 |
| 35 | TATTTAAAAG ATGCTGATGT CATTGTCAGT GCAGTTGGTA AGCCTGGTTT AGTAACAAAA | 7320 |
| | GATGTGGTCA AAGAAGGAGC AGTAATTATC GATGTTGGCA ATACGCCAGA TGAAAATGGC | 7380 |
| | AAATTAAAAG GTGACGTTGA TTATGATGCG GTTAAAGAAA TTGCTGGAGC TATTACACCA | 7440 |
| 10 | GTTCCTGGTG GCGTTGGTCC ATTAACAATT ACTATGGTAT TAAATAATAC TTTGCTTGCA | 7500 |
| | GAAAAAATGC GTCGAGGTAT TGATTCGTAA AGAGCCTGAG ACATAAATCA ATGTTCTATG | 7560 |
| | CTCTACAAAG TTATAATGGC AGTAGTTGAC TGAACGAAAA TTCGCTTGTA ACAAGCTTTT | 7620 |
| 15 | TTCAATTCTA GTCAACCTTG CCGGGGTGGG ACGACGAAAT AAATTTTACG AAAATATCAT | 7680 |
| | TTCTGTCCCA CTCCCTAATA ACTGAGTTTT AATGAAGTCT TTTAACCCAC ATTAAATATT | 7740 |
| - | ATTTTGCAAT TGCAATGAAT AACAAGAAAA ATCTGGGACA TTAATCGATC AAATGCTCCC | 7800 |
| 60 | TTCAAAGTAG ACATTGAATA AATGAAGGCT TTGAAGGGAG CATTTCACTT TGTACTTGGC | 7860 |
| | TCAACAATTT TATATAGACA GTAGTTAATT GAATGAAAAT AAGCTTGTAA CAAGTTTTCA | 7920 |

| | GTTGGGGATG | GGCGCCAACA | CAGAAGCTGT | GACTATGATA | AAGTACTACT | ACATAGTTAA | 8040 |
|-----|------------|------------|------------|------------|------------|--------------|------|
| | TCATTAGTGG | TTCTTTATCA | TTTTCGCCTC | CCTTTTCTTA | TTGTTTTGAT | ACACAAAAAT | 8100 |
| 5 | TTAAGTTCAA | ACTGTCGAAT | AAAGTTATAT | TTGATTTCAA | ATTATCCCTA | AATTATTAAT | 8160 |
| | TKTACAATTG | TGGCAGATTT | TCAAAATAAT | AATTATTTCC | TCATTATTTA | TAAATTTATA | 8220 |
| 0 | TTTAAATTTC | ATTCTTTATA | GGGTAAGATT | AGGACTATAG | TATGATGTGT | Arataatata | 8280 |
| · . | AATTAAGGTA | TAGTAAAGCT | AACTCAGAAA | TGACTTATCA | TTCGGAGGTT | ACATTATGAA | 8340 |
| | TAAACTATTA | CAGTCATTAT | CAGCCCTCGG | TGTTTCTGCT | ACACTAGTAA | CACCAAATTT | 8400 |
| 5 | AAATGCAGAT | GCAACGACGA | ATACTACACC | ACAAATTAAA | GGCGCTAATG | ATATCGTTAT | 8460 |
| | TAAGAAAGGT | CAAGATTATA | ACCTTCTAAA | CGGCATAAGT | GCATTTGATA | AAGAAGATGG | 8520 |
| | AGATTTAACC | GATAAAATTA | AAGTCGATGG | CCAAATTGAT | ACATCTAAAT | CTGGTAAATA | 8580 |
| 20 | TCAAATTAAA | TATCATGTCA | CTGATTCAGA | TGGTGCAATT | AAAATTTCCA | CTAGGTATAT | 8640 |
| | TGAGGTTAAA | TAGCCCTCAT | CACTATACTG | CAAATAAAAT | GGTAGCAAAC | GAACATGTTT | 8700 |
| | TGCTACCATT | TTATTTGTTA | TTCTAACTTC | ATCTGCAACT | TTAACCCAAA | TATTGTATTT | 8760 |
| 25 | TITCIGTATA | CCAAAGGACT | ACCTATCAAA | TTATTAAAAC | TTAACTGCTC | TTTTTAAAAA | 8820 |
| | AATGTTTTGA | TTTTGAACAA | ACAAATTTCC | ACTITICATT | GTTTAACGAT | AAATTACTTT | 8880 |
| | TGGCAAATTC | CTTATTAAAA | TGTTTGCGCT | TCCTTTCAAT | CAACTAGCCA | TCATTTTCAA | 8940 |
| 30 | TTTATTAGAC | AATTTCAAAC | TTTTTTTATT | TTCATTCAAT | TAACCTTTAA | TTGAAAGCTA | 9000 |
| | TTCTCAACTT | TCCTTTTAAA | TATGAAGCAA | TTTTTTCAAA | AACGCTATTA | GTCACAAAAT · | 9060 |
| 25 | GT ` | | | | | | 9062 |

(2) INFORMATION FOR SEQ ID NO: 86:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2738 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 86:

AAATATTTT TCAAAACTAT GTGAAAATGG ACCATGTCLA AATCATGTAA TAATGCAGYA 60
CATAATGCCA ACGGTCTmTC TTTATTGTCC CATGCATCAT GACCAATAAA TGACTCATCA 120
ATTAATCGTC TAACTATTTC ATACACACCT AAAGAATGTC CAAAGCGACT ATGTTCTGCT 180
GTGTGAAAAG ATAGGTACAG TGTTCCTAGT TGTCTAATTC GACGTAACCT TTGGAATTCC 240

55

40

45

| | | _ |
|-------------|---|------|
| | TCTTTAAAAA CTTTTTCTTC TACTAATTTT AAATCTACAT ATGCGTTAGT CATTATTCCC | 36 |
| | CTCCTTTTCG TITAATATAA TATTTAATTT ACTTAAAATG CTTTGTACAT AAGTGCTAAG | 420 |
| 5 | TCTAACTTTT CGCCATACAT TTCTGGCTCA TAAGAGCGTA AGATTGTAAA ACCTTGCTCT | 480 |
| | TTATAGTAAG CTACTGCTTC TTCATTTTTA TTATCTACTT CTAAGTAAAC ACCTTCAAAT | 54(|
| 10 | TTATCTTCAA AACGTGATAA TCCTTCATTT AACAATGCTG TACCATAACC TGTATGTTGC | 600 |
| 70 | GATTCTGGTT TAACATAATG AGCTGATAAA TATAATTCTT CACCGTAAAT AAAGTTAGCA | 660 |
| | AAGCCAACGA TGTCATTACC TTCTTCAACG ACTAAGAATA ATTGTTCTTG AAGTCTTTTC | 720 |
| 15 | TTTAAATGAT GTTCATTATA TGAAGCTECT AACAAGTGAT TAACTGTTGT CGCAGCGTAT | 780 |
| | ATATTTAAGT ATGTATTAAA CCAAGCTTTA GTTGCGACAT CTCTAATTTG AACAACATCT | 840 |
| | TTTTCAGTTG CTTGTCTTAC CTTGAACATG ACTTTCTCCC CTTATTAACA AGTTTTAATA | 900 |
| 20 | ACGGCATTAT ACCACAACTT GCTCAATACT TAATAAACAA TGATTGTCTA TTCAATTTAT | 960 |
| | ATATETATAT TTTCCGTTAA AATTAAAAAT AAAAAATAAC GAAGCAAAAA AECACTTCGT | 1020 |
| | TTAGTATGAG GTATGTCTTA TTGCAATATA CTATTCCACT CAGTTGCACG TGCTAAGGCA | 1080 |
| 25 | TAGTTGTCTT TCATGATGTC ACCAGGCTTT TCAGCAGTTC CAATAATATA ACCATTTAAA | 1140 |
| | GTGGCACCTA PAAAGTCTAA ACTATATTTC ATTTGCGTAA TTGCTGGTTC GCTTTTATTT | 1200 |
| | TTGGACAATC TCCACCAACT AAAATAACTC TAAAATCCTT TTCGGCCATT TGTGCCTTAA | 1260 |
| <i>30</i> · | AATTAGGATA TCGTTTATCT TGTAATGTTT CTGACCAATG TTCGATAAAT GCTTTCAATG | 1320 |
| | GTGCTGAAAT GCTATACCAA TACACTGGTG ATGCAAAAAT AATTGTATCA CTAGCCAATA | 1380 |
| | TTTTATCTAG AATCGGCAAA TAGTCATCGT CATATGAAGT AATAGTCTCT GCTGTATGTC | 1440 |
| 35 | TCACGTCACG TATCGGTTTA AACTGATGTT GTGTCACGTC AATCCATTGA TACTCTAAAT | 1500 |
| | CTTGCAAAGC GAATTTTGTT AATTGTGCAG TATTACCGTT TGGTCTACTC CCACCAAACA | 1560 |
| 40 | AAACAGTAAT CATTTTAGCC TAACCTCACT TTTGATTAAT AAATATCTGT GTTTTTCGTT | 1620 |
| +0 | ACCTAATTAT ACTATCATAA GCTTTGCCTA CCGAATAGTA AAACGCTTAC AACTTTTATA | 1680 |
| | TAAATTTGAC GAAATTTCGT CATGCCTTAT ATAACGTCGT TTGTGATACG GGGCTAATTC | 1740 |
| 15 | ATGATGAAAT TAGATACATA TATCACCATT AAATACAATT CATTTAGTCT TCAATCGGAA | 1800 |
| | ACAGTTCATC GATATATTGA ATCTCATCAT CTGATAAAAC GATATCTGCA GCTTTAATAT | 1860 |
| | TTTCAACGAC TTGTTCTGCA CGTTTTGCAC CAGGAATAAT CACATCGATA GCTGGTCTCG | 1920 |
| 50 | TTAAATAAAA TGCTAATACA ATGTTCGCAA TTGAAGTTTG ATGTGCTGCA GCTATGCTTT | 1980 |
| | CCAAAGCTIT TACGCGACGC ACATTITCIT CAAATACACC TGGTTTAAAA TCACGACGTG | |

| | GCTAATGGGA | AATATGGAAT | AAATGTGATT | TGGTGATCAA | CACAATATTG | TAATACTGCC | 2160 |
|----|------------|------------|------------|------------|------------|------------|------|
| | TCATTTTCGC | GATGCAATAA | ATTATATTCT | AACTGTACAA | CATCAACGTA | ACCATCTTTA | 2220 |
| 5 | TTTGCTTCTT | TAAGTTGATC | TAATGTGAAA | TTTGATACAC | CAATTGCTTT | AATCTTCCCT | 2280 |
| | TGTTCCTTAA | GCTCTTGTAA | TGCTGCAACT | GCTTGATCTT | TCGGAGTGTT | GTTATCCGGA | 2340 |
| | AAATGAATAT | AATATAAATC | GATATAATCA | GTTTGTAGAC | GTTTCAAACT | ATTCTCAACT | 2400 |
| 10 | TGTTGTTTTA | AATATTCCGG | TTGATTGTTC | TGATGTACTT | CTTGATTTTC | ATCAAATTCA | 2460 |
| | TGAGACCCTT | TCGTAGCAAT | TTTAATTTGC | TCTCGCGGAT | ATTCTTTAAC | AACTTCTCCA | 2520 |
| 15 | ACCAATTCTT | CTGATCGTTC | TGGCCCATAA | ATATATGCCG | татстаатаа | ATTAATACCA | 2580 |
| | TGATTAATGG | CTTGACGAAC | AACATCTTTT | CCTTGTTCTT | CATCTAAGTT | CGGATATAAA | 2640 |
| | TTATGCCCAa | CCTAtGCGTT | CGTCCCAAGT | GCGATTGGAA | ACACTTCAAC | ATCAGATTTA | 2700 |
| 20 | CCTAAGTTTA | CAAATTGCTn | CATTAGACCC | AGCnCCTT | | | 2738 |
| | | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 87:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9425 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 87:

| GATTAGATGA | TATTTAACGA | AAATTAaGrT | GMAATACTEG | AATGTATGAa | GTCTGATGTC | 60 |
|------------|------------|------------|------------|------------|------------|-----|
| GAAAATAGCT | ATTAAAATAG | AGTAGACGTA | ATGLAAATGA | AAGCACCTAA | AATAGAAAAA | 120 |
| TTTCAAAAAT | AGCGTAATTA | ТТАТААТААА | TAGACTGCCA | ATAAAATGCA | ATTTTTCACT | 180 |
| TATAÁCATTC | TTCAAAAAAT | AATAGCAAAA | TTATGTAAAA | AATATCTTGT | CATGGCAAGA | 240 |
| TTGGCTGTGC | TATAATCTAT | CTTGTGCTTA | AGAACGGCTC | CTTGGTCAAG | CGGTTAAGAC | 300 |
| ACCGCCCTTT | CACGGCGGTA | ACACGGGTTC | GAGTCCCGTA | GGAGTCACCA | TTTTTTAGGT | 360 |
| CTCGTAGTGT | AGCGGTTAAC | ACGCCTGCCT | GTCACGCAGG | AGATCGCGGG | TTCGATTCCC | 420 |
| GTCGAGACCG | TACAAATGCC | TATCCAAGAG | GATAGGCATT | TTTTTGCGTT | TAATATTATA | 480 |
| TTAATAAAAG | ATATATGGAC | GAATGATAAT | CATATTGATT | TATCTGTTCG | TCCATTTTCT | 540 |
| TTAAAATGTA | TGAACCTCAA | GTAACTTAGT | GGTTGGATAT | GAAAGATAAA | CGTAGACAAT | 600 |
| AAAATCTTTA | TTAGACGTAC | AAACATATGC | TACTGTCAAC | ATATTTCTTC | GTTGTGATAT | 660 |
| GCCACCAGTC | CTCCATAACA | TCAATTGTTA | AAGTAACGAA | TAACGAATAA | TGATATTTAT | 720 |

55

50

25

35

40

| | GACCTCATCA | TTGTGTTAAA | TATCATTGTC | ACAATCCGCC | GTGAGAAACT | TAAAAAAAA | 840 |
|-----------|------------|------------|------------|------------|-----------------|------------|------|
| | AGTAATATAT | AAGTTTATAT | TGGAAAATAG | AATTAATAGO | TTATAAATGG | TAAATTATAT | 900 |
| 5 | AATAGGTTAC | TATACGTTAT | AAGACGGAAA | ATGCGCACAA | TAACAAAAAT | AGTAAGCGAC | 960 |
| | ATCCTGTGAT | TTTTTACACA | AACATAAACG | ATAAAGAACA | AAAAATGATA | AAATAATATT | 1020 |
| | AATGATTTAA | GAAAAGAGGT | TTATGCAAAT | GGCTAGAAAA | GTTGTTGTAG | TTGATGATGA | 1080 |
| 10 | AAAACCGATT | GCTGATATTT | TAGAATTTAA | СТТАААААА | GAAGGATACG | ATGTGTACTG | 1140 |
| | TGCATACGAT | GGTAATGATG | CAGTCGACTT | AATTTATGAA | GAAGAACCAG | ACATCGTATT | 1200 |
| 15 | ACTAGATATC | ATGTTACCTG | GTCGTGATGG | TATGGAAGTA | TGTCGTGAAG | TGCGCAAAAA | 1260 |
| | ATACGAAATG | CCAATAATAA | TGCTTACTGC | TAAAGATTCA | GAAATTGATA | AAGTGCTTGG | 1320 |
| | TTTAGAACTA | GGTGCAGATG | ACTATGTAAC | GAAACCGTTT | AGTACGCGTG | AATTAATCGC | 1380 |
| 20 | ACGTGTGAAA | GCGAACTTAC | GTCGTCATTA | CTCACAACCA | GCACAAGACA | CTGGAAATGT | 1440 |
| | AACGAATGAA | ATCACAATTA | AAGATATTGT | GATTTATCCA | GACGCATATT | СТАТТААААА | 1500 |
| | ACGTGGCGAA | GATATTGAAT | TAACACATCG | TGAATTTGAA | TTGTTCCATT | ATTTATCAAA | 1560 |
| 25 | ACATATGGGA | CAAGTAATGA | CACGTGAACA | TTTATTACAA | ACAGTATGGG | GCTATGATTA | 1620 |
| | CTTTGGCGAT | GTACGTACGG | TCGATGTAAC | GATTCGTCGT | TTACGTGAAA | AGATTGAAGA | 1680 |
| | TGATCCGTCA | CATCCTGAAT | ATATTGTGAC | GCGTAGAGGC | GTTGGATATT | TCCTCCAACA | 1740 |
| 30 | ACATGAGTAG | AGGTCGAAAC | GAATGAAGTG | GCTAAAACAA | CTACAATCCC | TTCATACTAA | 1800 |
| | ATTTGTAATT | GTTTATGTAT | TACTGATTAT | CATTGGTATG | CAAATTATCG | GGTTATATTT | 1860 |
| | TACAAATAAC | CTTGAAAAAG | AGCTGCTTGA | TAATTTTAAG | AAGAATATTA | CGCAGTACGC | 1920 |
| <i>35</i> | GAAACAATTA | GAAATTAGTA | TTGAAAAAGT | ATATGACGAA | AAGGGCTCCG | TAAATGCACA | 1980 |
| | AAAAGATATT | CAAAATTTAT | TAAGTGAGTA | TGCCAACCGT | CAAGAAATTG | GAGAAATTCG | 2040 |
| 40 | TTTTATAGAT | AAAGACCAAA | TTATTATTGC | GACGACGAAG | CAGTCTAACC | GTAGTCTAAT | 2100 |
| | CAATCAAAAA | GCGAATGATA | GTTCTGTCCA | AAAAGCACTA | TCACTAGGAC | AATCAAACGA | 2160 |
| | TCATTTAATT | TTAAAAGATT | ATGGCGGTGG | TAAGGACCGT | GTCTGGGTAT | ATAATATCCC | 2220 |
| 45 | AGTTAAAGTC | GATAAAAAGG | TAATTGGTAA | TATTTATATC | GAATCAAAAA | TTAATGACGT | 2280 |
| | TTATAACCAA | TTAAATAATA | TAAATCAAAT | ATTCATTGTT | GGTACAGCTA | TTTCATTATT | 2340 |
| | AATgCACAGT | CATCCTAGGA | TTCTTTATAG | CGCGAACGAT | TACCAAACCA | ATCACCGATA | 2400 |
| 50 | TGCGTAACCA | GACGGTCGAA | ATGTCCaGAG | GTAACTATAC | GCAACGTGTG | AAGATTTATG | 2460 |
| | GTAATGATGA | AATTGGCGAA | ттасстттас | <u> </u> | CHIPCHICIDA A A | ~~~~~~~~~ | |

| | GTGATGGTAT | TATTGCAACA | GACCGCCGTG | GACGTATTCG | TATCGTCAAT | GATATGGCAC | 2640 |
|----|------------|--------------|-----------------|------------|------------------|---------------|------|
| | TCAAGATGCT | TGGTATGGCG | AAAGAAGACA | TCATCGGATA | TTACATGTTA | AGTGTATTAA | 2700 |
| 5 | GTCTTGAAGA | TGAATTTAAA | CTGGAAGAAA | TTCAAGAGAA | TAATGATAGT | TTCTTATTAG | 2760 |
| | ATTTAAATGA | AGAAGAAGGT | CTAATCGCAC | GTGTTAACTT | TAGTACGATT | GTGCAGGAAA | 2820 |
| 10 | CAGGATTTGT | AACTGGTTAT | ATCGCTGTGT | TACATGACGT | AACTGAACAA | CAACAAGTTG | 2880 |
| | AACGTGAGCG | TCGTGAATTT | GTTGCCAATG | TATCACATGA | GTTACGTACA | CCTTTAACTT | 2940 |
| | CTATGAATAG | TTACATTGAA | GCACTTGAAG | AAGGTGCATG | GAAAGATGAG | GAACTTGCGC | 3000 |
| 15 | CACAATTTTT | ATCTGTTACC | CGTGAAGAAA | CAGAACGAAT | GATTCGACTG | GTCAATGACT | 3060 |
| | TGCTACAGTT | ATCTAAAATG | GATAATGAGT | CTGATCAAAT | CAACAAAGAA | ATTATCGACT | 3120 |
| | TTAACATGTT | CATTAATAAA | ATTATTAATC | GACATGAAAT | GTCTGCGAAA | GATACAACAT | 3180 |
| 20 | TTATTCGAGA | TATTCCGAAA | AAGACGATTT | TCACAGAATT | TGATCCTGAT | AAAATGACGC | 3240 |
| | AAGTATTTGA | TAATGTCATT | ACAAATGCGA | TGAAATATTC | TAGAGGCGAT | AAACGTGTCG | 3300 |
| | AGTTCCACGT | GAAACAAAAT | CCACTTTATA | ATCGAATGAC | GATTCGTATT | AAAGATAATG | 3360 |
| ?5 | GCATTGGTAT | TCCTATCAAT | AAAGTCGATA | AGATATTCGA | CCGATTCTAT | CGTGTAGATA | 3420 |
| | AGGCACGTAC | GCGTAAAATG | GGTGGTACTG | GATTAGGACT | AGCCATTTCG | AAAGAGATTG | 3480 |
| _ | TGGAAGCGCA | CAATGGTCGT | ATTTGGGCAA | ACAGTGTAGA | AGGTCAAGGT | ACATCTATCT | 3540 |
| 30 | TTATCACACT | TCCATGTGAA | GTCATTGAAG | ACGGTGATTG | GGATGAATAA | TAAGGAGCAT | 3600 |
| | ATTAAATCTG | TCATTTTAGC. | ACTACTCGTC | TTGATGAGTG | TCGTATTGAC | ATATATGGTA | 3660 |
| 35 | TGGAACTTTT | CTCCTGATAT | TGCAAATGTC | GACAATACAG | atagtaagaa | GAGTGAAACG | 3720 |
| | raacctttaa | CGACACCTAT | GACAGCCAAA | ATGGATACAA | CTATTACGCC | ATTTCAGATT | 3780 |
| | ATTÇATTCGA | AAAATGATCA | TCCAGAAGGA | ACGATTGCGA | CGGTATCTAA | TGTGAATAAA | 3840 |
| 10 | CTGACGAAAC | CTTTGAAAAA | TAAAGAAGTG | AAGTCCGTGG | AACATGTTCG | TCGTGATCAT | 3900 |
| | AACTTGATGA | TTCCTGATTT | GAACAGTGAT | TATTATATTT | TCGATTTTAC | GTATGATTTA | 3960 |
| | CCGTTATCAA | CATATCTTGG | TCAAGTACTG | AACATGAATG | CGAAAGTACC | AAATCATTTC | 4020 |
| 15 | AATTTCAATC | GTTTGGTCAT | AGATCATGAT | GCTGATGATA | ATATCGTGCT | TTATGCTATA | 4080 |
| | AGCAAAGATC | GCCACGATTA | CGTAAAATTA | ACAACTACAA | CGAAAAATGA | TCATTTTTTA | 4140 |
| | GATGCATTAG | CAGCAGTGAA | AAAAGATATG | CAACCATACA | CAGATATCAT | CACAAACAAA | 4200 |
| 50 | GATACAATTG | ATCGTACGAC | GCATGTTTTT | GCACCAAGTA | AACCTGAAAA | GTTAAAAACA | 4260 |
| | TATCCCATCC | TATETA & CAC | C N TOTAL COTTO | CACAAAAMCA | A TOCOTTA TO COT | A TOTAL COLOR | 4330 |

| | GCAAACTATA | ACGATAAAAA | TGAAAAATAT | CATTATAAAA | ACCTGTCCGA | AGATGAAGCG | 444 |
|----|------------|------------|------------|------------|------------|------------|---------|
| | AGTTCCAGCA | AAATGGAAGA | AACGATTCCA | GGAACCTTTG | ATTTTATTAA | TGGTCATGGT | 4500 |
| 5 | GGTTTCTTAA | ACGAAGACTT | TAGATTGTTT | AGTACGAATA | ATCAGTCAGG | CGAGTTAACA | 4560 |
| | TATCaACGTT | TCCtTAATGG | TTATCCAACG | TTTAATAAAG | AAGGTTCTAA | TCAAATTCAA | 4620 |
| | GTCACTTGGG | GTGAAAAAGG | CGTCTTTGAC | TATCGTCGTT | CGTTATTACG | CACCGACGTT | 4680 |
| 10 | GTTTTAAATA | GTGAGGATAA | TAAATCGTTG | CCGAAATTAG | AGTCTGTACG | TTCAAGCTTA | 4740 |
| | GCGAACAATA | GTGATATTAA | TTTTGAAAAA | GTAACAAACA | TCGCTATCGG | TTACGAAATG | 4800 |
| 15 | CAGGATAATT | CAGATCATAA | TCACATTGAA | GTGCAGATTA | ACAGTGAACT | CGTACCGCGT | 4860 |
| | TGGTATGTAG | AATATGATGG | CGAATGGTAT | GTTTATAACG | ATGGGaGGCT | TGaATAAATG | 4920 |
| | AACTGGaAAC | TGACAAAGAC | ACTITICATI | TTCGTGTTTA | TTCTTGTCAA | CATCGTGTTA | 4980 |
| 20 | GTATCGATTT | ATGTTAATAA | AGTCAATCGC | TCACACATTA | ATGAAGTCGA | GAGTAACAAT | 5040 |
| | GAAGTTAATT | TTCAGCAAGA | AGAAATTAAA | GTACCGACTA | GTATATTGAA | TAAATCAGTT | 510 |
| | AAAGGTATAA | AATTAGAGCA | AATTACAGGG | CGATCAAAAG | ACTITAGTTC | TAAAGCTAAA | 5160 |
| 25 | GGCGATTCGG | ATTTGACCAC | ATCAGATGGT | GGAAAATTAT | TGAATGCGAA | CATTAGTCAA | 5220 |
| | TCGGTAAAGG | TCAGTGACAA | TAACTTAAAA | GATTTGAAAG | attatgttaa | CAAGCGCGTA | 5280 |
| | TTTAAAGGTG | CTGAATATCA | ATTAAGCGAG | ATTAGTTCAG | ATTCTGTAAA | ATATGAACAA | 5340 |
| 30 | ACGTATGATG | ATTTTCCGAT | TTTAAATAAC | AGTAAAGCGA | TGTTAAACTT | TAATATAGAA | 5400 |
| | GATAACAAAG | CGACTAGTTA | TAAACAATCA | ATGATGGATG | ACATTAAGCC | CACAGATGGT | 5460 |
| | GCAGATAAGA | AGCATCAAGT | GATTGGTGTG | AGAAAAGCAA | TCGAGGCATT | TATTATAT | 5520 |
| 35 | CGTTACTTGA | AAAAAGGTGA | TGAAGTCATT | AATGCTAGAC | TCGGTTACTA | CTCAGTCGTG | 5580 |
| | AATGĀAACGA | ATGTTCAATT | GTTACAACCA | AACTGGGAAA | TTAAAGTGAA | GCATGACGGT | 5640 |
| 10 | AAGGATAAAA | CGAATACTTA | CTATGTCGAA | GCGACAAATA | ATAACCCTAA | AATTATTAAT | .5.7.00 |
| | CATTAATATG | AATCGTAATA | AGCTAGCATT | GCAAGCTCAT | CATATGTGAG | AAGCGGTGCT | 5760 |
| | AGCTTTTTTG | CTGGTACGGT | TTATTATGGC | TGATGTTTTT | GCGTCTCCAA | CGTGCGCATT | 5820 |
| 15 | TATTCATATT | TTAAGTAGAA | CCGCATTGTA | AAATTAGTGT | AACTGTTATT | TTAAAAACTT | 5880 |
| | TAGTATTTGT | CTAATCATTG | TTATAATAAT | TAAGAAATTC | ATTGCACGTG | ATTATCAAAA | 5940 |
| | TTTAAATATA | AGAAACCGGT | CGATGAACTA | AAGTTACATA | ATAGGAAAGG | TATACAAAAC | 6000 |
| 50 | AGCTAATATA | CTGATAGTTT | CTGTAGGGAA | AATCGTATAT | TTGCACTGAT | GTATATTGCA | 6060 |
| | GTCATATAGA | GAGATTGACT | GTTTAAAGAG | ÄÄAGGATGAG | CCGCTTGATA | CCCATCACTC | 6120 |

| | TAGTTGATGT | TGGTTTGACT | GGAAAGAAAA | TGGAAGAATT | GTTTAGTCAA | ATTGACCGTA | 6240 |
|------------|------------|------------|------------|------------|------------|------------------|------|
| | ATATTCAAGA | TTTAAATGGT | ATTTTAGTAA | CCCATGAACA | TATTGATCAT | ATTAAAGGAT | 6300 |
| 5 | TAGGTGTTTT | GGCGCGTAAA | TATCAATTGC | CAATTTATGC | GAATGAAAAA | ACTTGGCAGG | 6360 |
| | CAATTGAAAA | GAAAGATAGT | CGCATCCCTA | TGGATCAGAA | ATTCATTTTT | AATCCTTATG | 6420 |
| | AAACAAAATC | TATTGCAGGT | TTCGATGTTG | AATCGTTTAA | CGTGTCACAT | GATGCAATAG | 6480 |
| 10 | ATCCGCAATT | TTATATTTTC | CATAATAACT | ATAAGAAGTT | TACGATTTTA | ACGGATACGG | 6540 |
| | GTTACGTGTC | TGATCGTATG | AAAGGTATGA | TACGTGGCAG | CGATGCGTTT | ATTTTTGAGA | 6600 |
| 15 | GTAATCATGA | CGTCGATATG | TTGAGAATGT | GTCGTTATCC | ATGGAAGACG | AAACAACGTA | 6660 |
| | TTTTAGGCGA | TATGGGTCAT | GTATCTAATG | AGGATGCGGC | TCATGCAATG | ACAGACGTGA | 6720 |
| | TTACAGGTAA | CACGAAACGT | ATTTACCTAT | CGCATTTATC | ACAAGACAAT | AACATGAAAG | 6780 |
| 20 | ATTTGGCGCG | TATGAGTGTT | GGCCAAGTAT | TGAACGAACA | CGATATTGAT | ACGGAAAAAG | 6840 |
| | AAGTATTGCT | ATGTGATACG | GATAAAGCTA | TTCCAACGCC | AATATATACA | ATATAAATGA | 6900 |
| | GAGTCATCCG | ATAAAGTTCC | GCATTGCTGT | GAGACGACTT | TATCGGGTGC | TTTTTTATGT | 6960 |
| ?5 | TGTTGGTGGG | AAATGGCTGT | TGTTGAGTTG | AATCGGCTTG | ATTGAAATGT | GTAAAATAAT | 7020 |
| | TCGATATTAA | ATGTAATTTA | TAAATAATTT | ACATAAAATC | AATCATTTTA | ATATAAGGAT | 7080 |
| | TATGATAATA | TATTGGTGTA | TGACAGTTAA | TGGAGGGAAC | GAAATGAAAG | CTTTATTACT | 7140 |
| 30 | TAAAACAAGT | GTATGGCTCG | TTTTGCTTTT | TAGTGTAATG | GGATTATGGC | AAGTCTCGAA | 7200 |
| | CGCGGCTGAG | CAGCATACAC | CAATGAAAGC | ACATGCAGTA | ACAACGATAG | ACAAAGCAAC | 7260 |
| | AACAGATAAG | CAACAAGTAC | CGCCAACAAA | GGAAGCGGCT | CATCATTCTG | GCAAAGAAGC | 7320 |
| 35 | GGCAACCAAC | GTATCAGCAT | CAGCGCAGGG | AACAGCTGAT | GATACAAACA | GCAAAGTAAC | 7380 |
| | ATCGÃACGCA | CCATCTAACA | AACCATCTAC | AGTAGTTTCA | ACAAAAGTAA | ACGAAACACG | 7440 |
| 10 | CGACGTAGAT | ACACAACAAG | CCTCAACACA | AAAACCAACT | CACACAGCAA | CGTTCAAATT | 7500 |
| | ATCAAATGCT | AAAACAGCAT | CACTTTCACC | ACGAATGTTT | GCTGCTAATG | CACCACAAAC | 7560 |
| | AACAACACAT | AAAATATTAC | ATACAAATGA | TATCCATGGC | CGACTAGCCG | AAGAAAAAGG | 7620 |
| 15 | GCGTGTCATC | GGTATGGCTA | AATTAAAAAC | AGTAAAAGAA | CAAGAAAAGC | CTGATTTAAT | 7680 |
| | GTTAGACGCA | GGAGACGCCT | TCCAAGGTTT | ACCACTTTCA | AACCAGTCTA | AAGGTGAAGA | 7740 |
| | AATGGCTAAA | GCAATGAATG | CAGTAGGTTA | TGATGCTATG | GCAGTCGGTA | ACCATGAATT | 7800 |
| 5 <i>0</i> | TGACTTTGGA | TACGATCAGT | TGAAAAAGTT | AGAGGGTATG | TTAGACTTCC | CGATGCTAAG | 7860 |
| | | | | | | 611611111 | 7020 |

| | TGAACCCATT AAACCCC | - |
|----------|---|-------|
| | TGAAGGCATT AAAGGCGTTG AATTTAGAGA TCCATTACAA AGTGTGACAG CGGAAATGAT | 804 |
| 5 | GCGTATTTAT AAAGACGTAG ATACATTTGT TGTTATATCA CATTTAGGAA TTGATCCTTC | 810 |
| • | AACACAAGAA ACATGGCGTG GTGATTACTT AGTGAAACAA TTAAGTCAAA ATCCACAATT | 8166 |
| | GAAGAAACGT ATTACAGTTA TTGATGGTCA TTCACATACA GTACTTCAAA ATGGTCAAAT | 8220 |
| 10 | TTATAACAAT GATGCATTGG CACAAACAGG TACAGCACTT GCGAATATCG GTAAGATTAC | 8280 |
| | ATTTAATTAT CGCAATGGAG AGGTATCGAA TATTAAACCG TCATTGATTA ATGTTAAAGA | 8340 |
| | CGTTGAAAAT GTAACACCGA ACAAAGCATT AGCTGAACAA ATTAATCAAG CTGATCAAAC | 8400 |
| 15 | ATTTAGAGCA CAAACTGCAG AGGTAATTAT TCCAAACAAT ACCATTGATT TCAAAGGAGA | 8460 |
| | AAGAGATGAC GTTAGAACGC GTGAAACAAA TTTAGGAAAC GCGATTGCAG ATGCTATGGA | 8520 |
| | AGCGTATGGC GTTAAGAATT TCTCTAAAAA GACTGACTTT GCCGTGACAA ATGGTGGAGG | 8580 |
| 20 | TATTCGTGCC TCTATCGCAA AAGGTAAGGT GACACGCTAT GATTTAATCT CAGTATTACC | |
| | ATTTGGAAAT ACGATTGCGC AAATTGATGT AAAAGGTTCA GACGTCTGGA CGGCTTTCGA | 8640 |
| | ACATAGTTTA GGCGCACCAA CAACACAAAA GGACGGTAAG ACAGTGTTAA CAGCGAATGG | 8700 |
| 25 | CGGTTTACTA CATATCTCTG ATTCAATCCG TGTTTACTAT GATATAAATA AACCGTCTGG | 8760 |
| | | 8820 |
| | CAAACGAATT AATGCTATTC AAATTTTAAA TAAAGAGACA GGTAAGTTTG AAAATATTGA | 8880 |
| 30 | TTTAAAACGT GTATATCACG TAACGATGAA TGACTTCACA GCATCAGGTG GCGACGGATA | 8940 |
| | TAGTATGTTC GGTGGTCCTA GAGAAGAAGG TATTTCATTA GATCAAGTAC TAGCAAGTTA | 9000 |
| | TTTAAAAACA GCTAACTTAG CTAAGTATGA TACGACAGAA CCACAACGTA TGTTATTAGG | 9060 |
| 35 | TAAACCAGCA GTAAGTGAAC AACCAGCTAA AGGACAACAA GGTAGCAAAG GTAGTAAGTC | 9120 |
| | TGGTAAAGAT ACACAACCAA TTGGTGACGA CAAAGTGATG GATCCAGCGA AAAAACCAGC | 9180 |
| | TCCAGGTAAA GTTGTATTGT TgtAGCGCAT AGAGGAACTG TTAGTAGCGG TACAGAAGGT | 9240 |
| 40 | TCTGGTCGCA CAATAGAAGG AGCTACTGTA TCAAGCAAGA GTGGGAAACA ATTGGCTAGA | 9300- |
| | ATGTCAGTGC CTAAAGGTAG CGCGCATGAG AAACAGTTAT TTCATAATCA ACAGTCATTG | 9360 |
| • | ACGTAGCTAA GTAATGATAA ATAATCATAA ATAAAATTAC AGATATTGAC AAAAAATAGT | 9420 |
| 45 | AAATA | 9425 |

(2) INFORMATION FOR SEQ ID NO: 88:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3886 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 88:

| | AGTTGTAATG | TCACATTTCC | AGAGTCTGAA | ATTATCTTTA | TCACGTTACA | TTTACTAGGC | 60 |
|------------|------------|------------|------------|--------------|-------------|------------|------|
| 5 | TCTAAAATGA | CTGAACATAC | AGCATCTTCA | ATTACCTTTG | AATACCATGA | TTTATCGCAA | 120 |
| | AATATACATG | AATTGATCAC | TTGTGTTAGC | CAAGAATTAG | GCATTGATAT | GTCAAAAGAC | 180 |
| | AACAAGTTAC | ATACCAGTCT | GATCACACAT | ATCAAACCAG | CTATACATCG | TATTAAATAC | 240 |
| 0 | GATATGCTAC | AACCTAATCC | TTTGAGGCAA | GAAGTTATGC | GTCGCTATCC | TCAAATCATT | 300 |
| | GAAGCCGTTA | GCAAGCATAT | TAGTCCAATT | GAACAAGATG | CTGCTATTCG | CTTCAACGAA | 360 |
| | GATGAATTAA | CATACATTAC | AATTCACTTC | GCATCAAGTA | TAGAGCGTGT | TGCAACACAT | 420 |
| 5 | AAACAATCAA | TGATTAAGGT | TGTCTTACTA | TGTGGTTCTG | GTATAGGCAC | GTCACAACTT | 480 |
| | ттаааатсаа | AACTAAATCA | CCTGTATCCT | GaGTTnCACA | TTTGGGAtGc | CTATTCCATT | 540 |
| 20 | TaTcAATTGG | aAGaAAGTCG | ATTATTGCAA | GATAACATTG | ATTATGTCAT | TTCAACAGTA | 600 |
| | CCTTGTGAAA | TATCAGCTGT | ACCAGTTATT | CATGTCGATC | CATTTATCAA | TCAACAATCT | 660 |
| | CGTCAAAAAT | TGAATCAAAT | TATCAATGAC | TCAAGAGAAC | AACGAGTCAT | GAAAATGGCA | 720 |
| 25 | ACTGATGGCA | AGTCACTCGC | AGATTTATTG | CCTGAACATC | GCATCATTAT | AAATAAACAA | 780 |
| | CCATTATCAA | TTGAATCCGC | AATTGCAGTG | GCTGTGCAAC | CTTTAATCAA | TGATGGCATT | 840 |
| | GTCTATTCAA | ATTATACAGC | TGCAATTTTA | AAACAATTTG | AACAATTCGG | GTCATATATG | 900 |
| 30 | GTCATTAGTC | CACATATTGC | ACTTATTCAC | GCTGGTACTG | ATTATGTACA | GAATGGTGTA | 960 |
| | GGTTTCGCAC | TAACATATTT | CACTGAAGGG | ATTATCTTTG | GTAGTAAAGC | TAACGATCCC | 1020 |
| | GTTCACCTTG | TAATTACATT | AGCAACGGAC | CACCCCAATG | CACATTTAAA | GGCATTGGGA | 1080 |
| 35 | CAGTTAAGCG | AATGCTTAAG | CAACGACTTA | TATCGACAAG | ATTTCTTAGA | TGGGAATATT | 1140 |
| | TTTAAAATTA | AACAACACAT | TGCTTTAACT | ATGACAAAGG | AGGCTTAATA | ACGTGTCATT | 1200 |
| | AGACATTTTG | TCAACAACAC | GCATCATTGT | AAAAGAACAA | GTAAATGATT | GGACTGAAGC | 1260 |
| 10 | TATAACTATA | GCTTCTCAGC | CATTACTACA | AGAACAAATT | ATTGAACAAG. | GCTATGTTCA | 1320 |
| | AGCAATGATT | GATAGCGTTA | ATGAACTTGG | ACCTTATATC | GTTATCGCAC | CTGAAATTGC | 1380 |
| 1 5 | AATTGCACAT | GCAAGACCGA | ACAATGACGT | ACATCAAGTT | GGTTTAAGTC | TATTAAAGTT | 1440 |
| | GAATCAACAT | GTGGCATTTT | GTGATGAAGA | TCACTACGCA | TCTCTCATTT | TTGTATTGAG | 1500 |
| | TGCCATCGAC | AATCATTCAC | ACTTATCTGT | ATTACAAAAT | TTAGCAACCG | TACTGGGCGA | 1560 |
| 5 <i>0</i> | TAACCAAACA | GTCCAGCAAC | TATTAACTGC | AACAAATGCA | CAAGACATTA | AAAACATTTT | 1620 |
| | ******* | C | *** | ma.cma.mcmcc | CCA CCCTTTA | CONNECTO | 1600 |

| | AAGTTGAACA 1 | CAGTGACATT | ATGACAGCAA | GTCCAGAGA' | I GGCTGACTTG | TTTATTTGTG | 1800 |
|------------|----------------|------------|------------|------------|--------------|--|------|
| _ | GTAGAGATTT A | GCTGAAAAT | GCCGAACGTC | TAGGGGATG | r cttagttctt | GATAATATTT | 1860 |
| 5 | TAGATAAAGC T | GAATTACAA | CAAAAGCTCT | CAGAAAAAT | T ACAACAACTT | AACATGATTT | 1920 |
| | AAAGGAGGTA C | | | | | | 1980 |
| 10 | AGCCATTCTT G | TTGCACTGA | TTGCCTTTAT | AGGTTTAATO | GTTCAGAAAA | AACCTGCCGC | 2040 |
| | AACGATCACT T | | | _ | | | 2100 |
| | TGATGTCGTC G | | | | | | 2160 |
| 15 | GCAAGGTATC G | | | | | | 2220 |
| | AACAGCTGCA C | | | | | | 2280 |
| | TAATTTAAAA T | | | | | | 2340 |
| 20 | AATCATTTTA AC | | | | | | 2400 |
| | ATTAGGATTA AT | | | | • | | 2460 |
| | AGGGAATGAC CA | | | | | and the second s | 2520 |
| 25 | GTAGGTCAAT TA | | | | _ | | 2580 |
| | TTAAGTTTCT TA | | | | | | 2640 |
| | ATCGCATGCT TA | | | | | | 2700 |
| 30 | TTTATTGTCT TT | - | | | | | 2760 |
| | ACGGGCGTTC GT | | | | | | 2820 |
| 35 | CTTGTACCAA AT | | • | | | | 2880 |
| ,5 | GCAGTATTAA TT | • | | | | | 2940 |
| | TTCTTATTTG GT | | | | | | 3000 |
| · o | • | | | | • | rggegeget | |
| | CTAAATGGTA TC | | | | | | 3120 |
| | GGTGGTGCTG CA | | | | | | 3180 |
| 5 | AACGCAGTAA AA | | | | | | 3240 |
| | ACAACAATTT TAT | | | | | | 3300 |
| | AAACACCGTC AC | | | | | | 3360 |
| 0 | GACGGTGTTT TGT | | | | | | 3420 |
| | | | | | | тсстраст | |

CCTAGACTCA AAATAAAGTC TGGTAATTTT TTAGTAGAAA CTTTTTGAGC TATTTCAGGT 3600
CTCTTTTCTT TAATTAATTT TGCAATTTCC AACAAATTAA TTTGTCCATC AGCCGTCGCA 3660
ATAAATCGCT TGCCATTAGC TTGTTCATTT GTCATTGCCA AAATGTGCAG TTCAGCTACG 3720
TCTCTCACAT CAACAACATT TAACGGAATT TGCGGTACAC GTTTCATTGA ACCATTCAAT 3780
AAATTTTCTA ATAAATGAAA GCTTCCTGAA ACGTGTGCAT CTAATGATGG CCCAAAAAATT 3840
GCAACTGGAT TGATTGTGGC AAATTCTACT GTTGTATTTT CATTCT 3886

(2) INFORMATION FOR SEQ ID NO: 89:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4879 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

20

25

30

35

40

45

50

5

10

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 89:

| . 60 | TTAATTTCCC | AATTAATAAT | GACAATTATT | TATACAGACC | AAAAATFIGG | GTCATCTATC |
|------|-------------|------------|------------|------------|------------|------------|
| 120 | TTAGAATCTA | AGAACAACCA | TACAACATAA | TATCCACAAA | AGTGATTAAA | AGGCAATACC |
| 180 | TTCTCAGTGG | GTGTTTAAAT | CTAATCAATA | AAGATGACAT | TATATTGTCT | TTTCACAACT |
| 240 | AATACGCAAA | GATACTTTAA | CGTAAACTTT | TACTATAAAA | GGTTTAAAAG | CTGTGAATGA |
| 300 | CTTGCATCAT | TATTTTTTAA | TTTACCGTTT | TATTATAGAG | CCCTAATTCA | AAACGGTAAA |
| 360 | CCAGTATAAT | CATTGCTTGT | GATCAGTAAC | TGGTAGTTTG | ACATTATTGT | AGTTATATTA |
| 420 | AACACATACT | TTTAAAACTA | CATTTGGGTC | TTTCCaTTTG | AATTGAATAT | CAACCGTTAC |
| 480 | TTACTACCAA | CACTTTATTA | AATCATTATA | TCAATAGAAT | ATTATGTTCT | TATAGTTGCC |
| 540 | AATTGCTGTA | GAAATTTGGC | AAGCTGTTTG | GCCGCATTTA | TTCATTATTA | ATTIATITGC |
| 600 | GCGGCATTTT | TGCAACCGGC | AAATTTGACG | TTAAACGGAT | TTTATTTCCA | AAGCTTGATT |
| 660 | TTATTACTTG. | ATTCGCTTGG | TTTGATTATT | TAACTTGATT | TGGTGCAACG | GnCCATAATA |
| 720 | GATTGTTGCT | ATTTGAATTA | TGTTTGTTGC | TTTTGGTCAT | ATTTGTTTGG | ATTGGTTATT |
| 780 | TCATCTTTAT | GTCTTTACTA | CTTTGTTTAC | TCTTTATTAT | TGCACTATTA | GGTTATCGTT |
| 840 | GTTTTCGCTT | TTGTTGTTCA | TTTTTTTATC | GAATCATTTG | ATCTTTAGAT | TATCTTTCTT |
| 900 | CATGCAGCTA | ATCTTGACCA | GTTGGTCACT | CCGTCTTTTT | TTCTTTATTA | TATCATCTTT |
| 960 | TCCTCCTATA | CATACATATC | CTAATCTTTT | AACCCTGTAA | TAATGCTAGT | AAAATAATGA |
| 1020 | TTTCATGGCT | ATGTGTATCT | CATATCTACC | TCTTGAAATA | CATTGAATAA | ATTCGATATT |

| | TAAGGTTCTT | TTTATTATAC | CCTAATTTT | GTTCATTATT | TTTAATTTA | TGTGAATTTT | 1140 |
|---------|------------|------------|------------|------------|------------|------------|------|
| | ATGETTRCTA | TAAATTTAAT | TATTTTACTI | TAACAATTCA | TTACGCATTT | AGCATTTCAA | 1200 |
| 5 | GGTATACACA | ATATTTATTA | CTATGATTTC | ATTTTATCTG | CTGCAAAAAC | AATCATTATA | 1260 |
| | ACTCTTTTTC | CATAATTAAA | TCTGTATCCG | TTACATCACC | TGTTTGAAAA | TGATGTTCAC | 1320 |
| | CAACCACTTT | AAATCCATGA | CGTTTATAAA | ATGCTTGAGC | ACGAGGATTA | TGCTCCCAAA | 1380 |
| 10 | CTCCTAGCCA | AATTTTATGT | TTATTATGTT | CTTGAGCAAT | TTTTTCGGCC | AATTCTATCA | 1440 |
| | ATTGTGAACC | TCTTCCGCCA | CCTTGAAAGT | CTTTCAAAAA | ATATATGCGC | TGCACTTCTA | 1500 |
| 15 | AATAGGTCTC | CCCCATTTCT | TCAGTTTGAG | CACTATTAAT | ATTCATCTTT | ATATAACCAA | 1560 |
| | CATTCGCACC | ATCTTCTTGa | TAAAAATAAT | GAAATGAATC | TACATGGTTA | ATCTCTTGTG | 1620 |
| | TAAATTTCTC | TACAGTATAA | TTGTCTTTAA | AAAATTGATC | AAAATCTTTG | TCATCATAGT | 1680 |
| 20 | AAGAACCAAA | CGTGTCATAA | AATGTTCTAG | TTGCTAATTC | AACTAATTCA | CTAGCATTTT | 1740 |
| | GTTCTGAAAT | TTCTTTGATT | ATCCCAGCCA | TATAAATCCT | CCAATAAACA | GTGATCGAAT | 1800 |
| | CAAAATATTA | CTTATGTTAT | TTTTCAGCCA | AAACTATTTA | AAAATACATT | AACACAAATC | 1860 |
| 25 | AATTACAAAT | TGTATTGATT | GTGTGTAACA | TCAATAAATG | ATACATTTAT | TCCAGTAAAA | 1920 |
| | TGGCCGTATT | TTCAAAAGAG | AAAAAGAGAG | GATGTATCGT | TGTGATAGAA | ACATTTAAAG | 1980 |
| | CGTTTGTAAT | TGATAAAGAT | GAGAGTGGTA | AAGTGACACC | AACTTTCAAA | CAATTATCGC | 2040 |
| 30 | CTACTGATTT | ACCTAAAGGA | GATGTGCTGA | TTAAAGTACA | TTACTCTGGT | ATAAATTATA | 2100 |
| | AAGATGCTTT | AGCGACTCAA | GATCATAATG | CAGTCGTAAA | ATCGTATCCT | ATGATTCCAG | 2160 |
| 35 | GAATAGATTT | AGCTGGAACA | ATTGTTGAAT | CCGAAGCACC | AGGCTTTGAa | AAAGGAGAAC | 2220 |
| | AAGTAATTGT | AACGAGTTAT | GACCTAGGTG | TCAGCCATTA | TGGCGGTTTT | AGTGAATATG | 2280 |
| | CGCGTGTAAA | ATCAGAATGG | ATTATCAAGC | TTCCTGATAC | TTTAACATTA | GAAGAATCAA | 2340 |
| 40 | TGATATATGG | CACAGCTGGT | TATACTGCCG | GTTTAGCAAT | TGAAAGACTT | GAAAAAGTTG | 2400 |
| | GAATGAATAT | TGAAGATGGT | CCTGTACTCG | TTCGCGGTGC | TTCAGGTGGT | GTCGGTACTT | 2460 |
| | TAGCAGTACT | CATGCTTAAT | GAACTTGGTT | ATAAAGTTAT | CGCAAGTACA | GGTAAACAAG | 2520 |
| 45 | ATGTTAGCGA | TCAATTACTT | GAACTTGGTG | CCAAAGAAGT | TATCGATCGA | CTTCCTGTTG | 2580 |
| | AAGATGATCA | TAAAAAGCCA | CTCGCATCAT | CAACTTGGCA | AGCTTGTGTA | GACCCTGTTG | 2640 |
| | GTGGCGAAGG | TATTAATTAT | GTTACAAAGC | GTTTAAATCA | TAGTGGGTCA | ATTACAGTTA | 2700 |
| 50 | TTGGTATGAC | TGCCGGTAAT | ACTTATACTA | ATTCTGTATT | CCCTCACATT | TTAAGAGGTG | 2760 |
| | TAAACATTTT | AGGAATTGAC | TCGGTATTTA | CTGCTATGAA | ATTAAGACAG | CGCGTTTGGC | 2920 |

| | TIGATGAACT | TCCAGAACAA | CTTAACAAAG | TAATTAAACA | TGAAAATAAA | GGGCGCATTG | 2940 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | TTATCGATTT | CGGTGTAGAT | AAATAGTATT | CATGAAAAAG | ACATCCCGTT | ATGCGAGATG | 3000 |
| 5 | TCTTTTTTAA | TTTAGTATTT | GATATACATA | CCGCCTGAAT | CTGGTTCGGT | AGGTATAAAT | 3060 |
| | CCAAATTTTG | TATATAATTT | ATCCGCTGGG | TAGTCTGCAA | TCAGACTAAC | GTATGTACTC | 3120 |
| | TCAACAGCCA | CACCTTTAAT | ATATTGCATA | ATATGCTCCA | TAATTAGACT | GCCGTAACCT | 3180 |
| 10 | TGACCTTGGT | AACTTTTCAA | AACTGCAATA | TCAACAATTT | GAAAAACAGT | TCCGCCATCG | 3240 |
| | CCAATCACTC | TACCCATACC | AATTAACCGA | TCTTTATCAT | ACAAGGTTAC | TGTAAATAAG | 3300 |
| 15 | GCATTAGGTA | ATCCTTTTTC | aGCTGTTCGC | GCGTCTTTGG | ACTCATACCT | GCGTTAATCC | 3360 |
| ,,, | TTAATGCGCA | ATAATCCTCG | CAAGTCGGAA | TATCATATGT | CACTTTAACC | ATTATTTACC | 3420 |
| | CCACTTTTCA | TCACACAATA | TATCAACCTA | GTATAAATGT | TTATTTACAA | TAGTCTTATT | 3480 |
| 20 | CGCTTCTTTA | AACACTTCAT | GATGACTTGA | AACATAACCC | TCTGCATTCG | CATCTGGTTG | 3540 |
| | GATATATGTT | TŤĀGCAAGGT | TCGCTGCATT | TGCACCATÇA | CTAAATGCAC | TTGCAATTAG | 3600 |
| | ATGTGATTTT | GCATCATGAT | AAACAATATC | TCCACACGCA | TAGATACCAG | GTATACTAGT | 3660 |
| 25 | TGTCGTATTA | CCAAATCCTT | TAACACGACA | ATCATCATGC | ATATCTAGCT | TTGAAGATGT | 3720 |
| | TECACTCAAT | AATGTATTAC | AACGATCAAA | CCCATGACTA | ATAATGACAT | CGTCAAATTT | 3780 |
| | AACTGTATGC | CTATCGCCAC | TTTCAACATG | TTCCAAAACA | ACTTCACTTA | TATGCGTTTC | 3840 |
| 30 | ATCATCATTG | CCGACCAAGT | ATTTAATACG | TGTTTTTGGG | CATAGTTTCA | CATTTAAATC | 3900 |
| | TGTCACCAAC | GTTTTCATCG | CTTCATGACC | ACTTACATCT | TCTTTTCGAT | AAACAACTGT | 3960 |
| | CACGCTTTTA | GCAATCTTGG | CAATATCATG | CGCCCAATCT | AATGCTGTAT | TTCCTCCACC | 4020 |
| 35 | TGATATTAAT | ACATCTTTAT | CTTTGAAACG | TCTGTAACTT | TGTACAACAT | AATGTAAATT | 4080 |
| | AGTTAATTGA | TATCTCTCTA | CACCITTAAC | ATCTAATTGT | TTTGGATTAA | TAATACCCGC | 4140 |
| 10 | ACCAATTGCA | ATGATAACTG | CTTTCGATGT | ATATATTTCT | CCCGCTTCTG | TTTCAACTTC | 4200 |
| •• | GAAATGACGT | TCTGCCTTTT | TCCTAATATC | TACCACACGT | TCATTCAAAT | GAACTTCCGG | 4260 |
| | TTTAAAATAT | AATCCTTGCT | TAATTGTATC | TTTTAAAATT | TCATGACAAG | GTTTTGGCGC | 4320 |
| 4 5 | AATGCCGCCA | ATATCCCAAA | TAATTTTTTC | AGGGTAAATT | CTCATCTTAC | CCCCTAATTC | 438 |
| | AGATTGAACA | TCTATCAATC | TTACAGACAT | ATCTCGCAAT | CCAGCATAAA | AGCTTGCATA | 444 |
| | CAAACCAGAC | GGACCGCCAC | CAATGATTGT | AACATCTTTC | ATTATGTGCC | TCCTATGACT | 450 |
| 50 | CTCTATATTC | ATTTCTTTCA | TTAACGTGCT | CAAATTGATA | ATTATTATCA | TTTAAAGCCA | 456 |
| | TTATACTATT | AATATTTATA | TTGTTAAAAT | AAATCGCATA | GTTAGCCATG | AATTATCAAT | 462 |

| | | _ |
|------------|--|------|
| | GAAAGATGTG TATATTTTT AGTTCTAGTT ATATTATTTT TTAAAAGACT CATCACGTGG | 4740 |
| | TTCTTTAAGA ATTGCTTGTC TTAAAAGGAA AAATAGCAAC AATAAACCTG CAAGCATACC | 4800 |
| 5 | TGTGTGCCCA ATACCTGCAA AGCCTGCNAA TGCTTCTGGA GAGTATGATT TACCAGTGAC | 4860 |
| | TTGGAAGAAT CCTTTTGTC | 4879 |
| 10 | (2) INFORMATION FOR SEQ ID NO: 90: | |
| 15 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1560 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 90: | |
| 20 | ATAATGTCTT AGATTGATTG GGAGTTTTTT TAATTTTTTT GAAATTAAAT TAATCTGTAS | 60 |
| | YTAATAAAA ATTTGAATAA CTGACACAYT TTTTTGATCA TAGCTAYATA CTTTGTGAAT | 120 |
| | TAATTCACAT TATAATAAGA GTGAAGATAA GAGTATTATA AATAATCTTT AAATAAATAT | 180 |
| 25 | ATGTGAAGTA AAAATTACAC GTTAGCATAT CGATTATGGT CATTTCKTTT AACATATTAA | 240 |
| | CTGGGGAACG TTAAAAGTTA ACGGKTGATA TCYAACLAAA AACAAGGTCA CAGTAGTATG | 300 |
| | TTTTAATCTG GCGTCTATTA CAAATAAAAA TTACATCTAT AATTATTCGT TTTCTTTTTT | 360 |
| 30 | GAAAGTAATA GCCAATTAAT ATCATACATA CTGGAGTGAC TATAAGGAGG ACATTATTAT | 420 |
| | GAGAGCAGCA GTTGTAACGA AAGATCACAA AGTAAGTATT GAGGACAAAA AGTTAAGAGC | 480 |
| 35 | TTTAAAACCT GGTGAAGCGT TGGTACAAAC GGAATATTGT GGCGTTTGTC ATACCGATTT | 540 |
| <i>3</i> 5 | ACATGTTAAG AATGCTGATT TTGGTGATGT TACAGGCGTT ACTTTAGGTC ATGAAGGTAT | 600 |
| | TGGTAAAGTC ATCGAAGTTG CGGAAGATGT AGAATCATTA AAAATTGGAG ACCGTGTGTC | 660 |
| 40 | TATCGCTTGG ATGTTCGAAA GCTGTGGAAG ATGTGAATAT TGTACAACAG GTCGTGAAAC | 720- |
| | ACTITICCCT ACTGTGAAAA ATGCTGGTTA TACAGTAGAT GGTGCAATGG CTGAACAAGT | 780 |
| | TATTGTTACT GCAGACTATG CTGTGAAAGT ACCTGAAAAA TTAGATCCAG CAGCAGCGTC | 840 |
| 45 | TTCTATTACA TGCGCAGGTG TGACAACTTA TAAAGCTGTA AAAGTAAGTA ATGTAAAACC | 900 |
| | TGGACAATGG TTAGGTGTTT TTGGTATAGG TGGTTTAGGT AACCTAGCTT TACAATATGC | 960 |
| | TAAAAACGTT ATGGGGGCTA AAATTGTTGC CTTCGACATC AATGATGATA AATTAGCATT | 1020 |
| 50 | CGCGAAAGAA TTAGGTGCTG ATGCTATTAT TAATTCTAAA GATGTTGATC CAGTTGCAGA | 1080 |

55

1140

AGTTATGAAA TTAACTGATA ACAAAGGATT AGATGCAACA GTGGTAACTT CAGTTGCTAA

TTTACCTGTT GATAAAATGA ACTTAGATAT CCCAAGATŁA GTGCTTGATG GTATTGAAGT

AGTAGGTTCA CTTGTTGGTA CAAGACAAGA CTTACGTGAA GCGTTTGAAT TTGCTGCTGA

AAATAAAGTA ACACCTAAAG TTCAATTAAG AAAATTAGAA GAAATCAATG ATATTTTTGA

AGAAATGGAA AATGGTACTA TAACTGGTAG AATGGTTATT AAATTTTAAA AATATCAACT

GACTATATAG ATAAAGAAGG TAGTGCTCTG AACACTATCA TTATTAATCA AACCCCGAGG

TTTTCCTGAA AAGATAGTGG NAAATCCCCG TGTTTTTTGG GTTTGAGGNG GTTGTNTGTA

(2) INFORMATION FOR SEO ID NO: 91:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11014 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

20

25

30

35

40

45

50

5

10

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:

GTCCTGTnGC TGCAATGAAT ACGCCTAAAA ATCCAGGGAT GTAATGGATA CTTTGTGGTA 60 GTACTAATGA TAGAAATGAT AAAAATGAAA TCACAAAGGC TACGCTCGCA AAAGCTTGAC 120 ATGTACGCTT ATCGCCATAA TCTAACCCTG TACGTATATG TAATAAATAC TGTAATCCGA 180 TACTTAAATA CATAATTGCC ACGCATAAGA AGAATGGGAA GAATGTCTTT TCAAAGTCCG 240 GATATAGGCT GTTAGATAGG AAGACCATGA TGAACATATT AAACATCATA AACGAGACGT 300 CTTTGAATGT AACTTGACCA AATCGATTTG TAAAAAATGT TTGATGAGAC CACATTAACC 360 ATAAGAACAA ACTCATGACG ATGTATTTGA AAAATAAATC AGCTGAAATG GAACCGTTTT 420 GTGTTGTTAA AATCACATGT GCAATTTTTT GAATGGCATA GACGAAAATT AAATCAAAGA 480 ACARCTCATG GAATCCTGCA CGCTTTTCAG CTAAATGTTT TGGTGTTAAT GCATTAACCA 540 TAAAATTITA ACTCCTITAA GATGTGTAAT TAATTTACTA AGTATACTAT TTATTTTTTC 600 TAGTGAATAG GGGCAGATTT GGCGATGAAG TGGAAGGAGA GGTGACTGCA AGGTAATTGC 660 GGAATTAACA ATCATCAGCG ATTTAATATT TGACTGGAGA CGTCATGGTA ATAAAAAATT 720 780 GATGAGAAAT TGATGGTGAA ACCAGCTGTG AATASCGATG CAATGATYSA TAGAATTTAA TTAGAGTCAT TACGCGAAAT GATTAATGAT AATTTGTGGT AAATCAAAGC ATAATTTTGT 840 900 ACTATAGATG AGGATGATAG AGCATATTTA AGAGGGTGAA ATGTTAAAGT GAAACCGTTT ACGTTTCCGA TTGCCCAAAC AAATTACATC ATTGTATAAT ATGATTTGTT AAATGCATAA 960 CAAGAATGAA AATGTAACAT ACGTAGCAAT TGGTTTCATA AATTGGATGT TAGTGGCGTA 1020

| | TGACGAGAGT | CGTATTAGCA | GCAGCATACA | GGACACCTAT | TGGCGTTTTT | GGAGGTGCGT | 114 |
|------------|------------|----------------|--------------------|------------|-------------|------------|-------|
| _ | TTAAAGACGT | GCCAGCCTAT | GATTTAGGTG | CGACTTTAAT | AGAACATATT | ATTAAAGAGA | 120 |
| <i>5</i> | CGGGTTTGAA | TCCAAGTGAG | ATTGATGAAG | TTATCATCGG | TAACGTACTA | CAAGCAGGAC | 126 |
| | AAGGACAAAA | TCCAGCACGA | ATTGCTGCTA | TGAAAGGTGG | CTTGCCAGAm | ACAGTACCTG | 132 |
| 10 | CATTTACGGT | Gaataaagta | TGTGGTTCTG | GGTTAAAGTC | GATTCAATTA | GCATATCAAT | 138 |
| | CTATTGTGAC | TGGTGAAAAT | GACATCGTGC | TAGCTGGCGG | TATGGAGAAT | ATGTCTCAAT | 144 |
| | CACCAATGCT | TGTCAACAAC | AGTCGCTTTG | GTTTTAAAAT | GGGACATCAA | TCAATGGTTG | 150 |
| 15 | ATAGCATGGT | ATATGATGGT | TTAACAGATG | TATTTAATCA | ATATCATATG | GGTATTACTG | 1560 |
| | CTGAAAATTT | AGTAGAGCAA | TATGGTATTT | CAAGAGAAGA | ACAAGATACA | TTTGCTGTAA | 1620 |
| | ACTCACAACA | AAAAGCAGTA | CGTGCACAGC | AAAATGGTGA | ATTTGATAGT | GAAATAGTTC | 1680 |
| 20 | CAGTATCGAT | TCCTCAACGT | AAAGGTGAAC | CAATCGTAGT | CACTAAGGAT | GAAGGTGTAC | 1740 |
| | GTGAAAATGT | ATCAGTCGAA | AAATTAAGTC | GATTAAGACC | AGCTTTCAAA | AAAGACGGTA | 1800 |
| | CAGTTACAGC | AGGTAATGCA | TCAGGAATCA | ATGATGGTGC | TGCGATGATG | TTAGTCATGT | 1860 |
| 25 | CAGAAGACAA | AGCTAAAGAA | TTAAATATCG | AACCATTGGC | AGTGCTTGAT | GGCTTTGGAA | 1920 |
| | GTCATGGTGT | AGATCCTTCT | ATTATGGGTA | TTGCACCAGT | TGGCGCTGTA | GAAAAGGCTT | 1980 |
| | TGAAACGTAG | TAAAAAAGAA | TTAAGCGATA | TTGATGTATT | TGAATTAAAT | GAAGCATTTG | 2040 |
| 3 0 | CAGCACAATC | ATTAGCTGTT | GATCGTGAAT | TAAAATTACC | TCCTGAAAAG | GTGAATGTTA | 2100 |
| | AAGGTGGCGC | TATTGCATTA | GGACATCCTA | TTGGTGCATC | TGGTGCTAGA | GTATTAGTGA | 2160 |
| 35 | CATTATTGCA | TCAACTGAAT | GATGAAGTTG | AAACTGGTTT | AACATCATTG | TGTATTGGTG | 2220 |
| | GCGGTCnAAC | TATCGCTGCA | GTTGTATCAA | AGTATAAATA | ATAAGAAAAC | AGGTTATCAC | 2280 |
| | AACAGTATTA | ATLACATGTT | GGCATAACCT | GTTTTTATTT | GTTTATGGAT | TTATTGGGTA | 2340 |
| 10 | ATATTAGTCA | TTTGATGGTT | TAATTGCAAA | TGCTCTAACA | GGGAACCCAG. | GTGCATCTTT | -2400 |
| | TGGTTTAGGG | CTGATAGCGT | AAATGATGGC | GCCACGAGTT | GGTAATTGAT | CTAAATTAGT | 2460 |
| | TAATAACTCG | ACTTGGTATT | TATCCTGACC | AAGAATATAA | CGTTCGCCAA | CTAAATCACC | 2520 |
| 15 | ATTTTTTACA | ACGTCCACAG | ATGCATCGGT | ATCGAATGTT | TCATGACCAA | CAGCTTCAAC | 2580 |
| | ACGACGTTCT | TCAATTAAGT | ACTTCAAAGC | ATCTAATCCC | CAACCCGGTG | CATGTTGTTG | 2640 |
| | TCCGTTCGCA | TCTTTGTTTT | CAAACTTTTC | AATATTAGGC | CAACGTTTTG | ACCAATCGGT | 2700 |
| 50 | ACGAAGTGCA | ACAAAAGTGC | CAGGTTCAAT | AGTACCATGC | TCTTTTTCCC | ATGCTTCTAT | 2760 |
| | ATGCGCACGT | מבידמבים מידים | የምርም የ ሳም የ | | | | |

| | AAAGTGAATT | GGTGCATCAA | TGTGAGTACC | ATATTGCGTT | ACAATATTCC | AACGTTGCAC | 2940 |
|----|------------|------------|------------|------------|------------|------------|------|
| | ATAGAAACCA | TGATCTTTAA | CCGTGAATAA | AGTTGAAACT | TCGCCTTTTT | CAAACTCACT | 3000 |
| 5 | AAAACGTGGT | ATTTCCGGAT | CAAATGTATG | CGTTAAATCA | ACCCAAGTTG | CTTGTTTTAA | 3060 |
| | AGTATTTAAT | TGTTGCCATA | AAGGATATTG | TGTCATAAAA | TCACCCGTTT | TTAGTTTATT | 3120 |
| 10 | ATATGATAAA | TGCTGCGATT | ATTCTTGGCG | TTTAGCTTTA | ACAGCATTCA | CAAGCACAGT | 3180 |
| | CAATGCATCT | TTAACTTCTT | CTTCTTTTCG | CGTTTTTAAA | CCACAGTCAG | GGTTTACCCA | 3240 |
| | GAATAATGAG | CGGTCGATTT | GTTGTAGTGA | ACGATTGATT | GCTGTAGTAA | TTTCTTCTTT | 3300 |
| 15 | TGTTGGAATA | CGTGGACTAT | GAATATCATA | TACACCTAGA | CCAATACCTA | AATCATAATT | 3360 |
| | AATATCTTCA | AAGTCTTTAA | TTAAATCACC | ATGGCTACGA | GATGTTTCAA | TTGAAATAAC | 3420 |
| | ATCAGCATCT | AAGTCATGAA | TAGCATGAAT | GATTTGACCG | AATTGAGAAT | AACACATATG | 3480 |
| 20 | TGTATGGATT | TGAGTTTCAT | CACGAACTGA | AGACGTTGCA | AGTTTAAATG | ATAAAACAGC | 3540 |
| | ATCTTTAAGA | TATTGTTCGT | GATATTCAGA | GCGTAATGGT | AAGCCTTCAC | GTAATGCAGG | 3600 |
| | TTCGTCAACT | TGGATAACTT | TGATTCCTGC | AGCTTCAAGT | GCTAATACTT | CTTCGTTGAT | 3660 |
| ?5 | TGCTAAAGCA | ATTTGATCTT | GAACGACTTT | ACGTGGTAAA | TCAACACGTT | CAAATGACCA | 3720 |
| | GTTTAGAATT | GTTACAGGTC | CAGTTAACAT | ACCTTTAACT | GGTTTATCTG | TTAAGCTTTG | 3780 |
| | TGCATAAACT | GTTTCATCAA | CAGTTAAAGG | CGCTGTCCAT | TTTACATCAC | CATAAATGAT | 3840 |
| 30 | TGGTGGTTTT | ACGGCACGTG | AACCATATGA | TTGCACCCAA | CCGAATTTAG | TTACTAAGAA | 3900 |
| | ACCTTGTAAT | TTTTCTCCGA | AGAATTCAAC | CATGTCATTA | CGTTCAAATT | CACCGTGAAC | 3960 |
| 35 | TAATACATCT | AAGCCAATGT | CTTCTTGAAT | TTTAATCCAT | CGAGCAATTT | CATTTTTTAA | 4020 |
| | GAATGTTTCA | TATGCTTCGT | CTGTAATGCG | TTTGTTCTTC | CAATCTGCAC | GGTATTTTCG | 4080 |
| - | AACTTCTCGG | CTTTGTGGGA | ATGATCCAAT | AGTTGTTGTT | GGTAAATCCG | GTAAGTTCAA | 4140 |
| 10 | ACGTTTTTGT | TGTTGTTCAA | TACGTTGCGC | GAATGGTGAT | TGTCTTGAAG | TACGCACGCT | 4200 |
| | TTCGAAATCA | TAATCTAAGT | TTTTGAATGA | TIGATITIGG | AAACGCTCAT | AACGTGCTTT | 4260 |
| | TAATTTATCA | TATTTAACAC | TATCGTTTTG | ATTAAATAGG | CGACGCAATG | CATCTAATTC | 4320 |
| 15 | GTCTAATTTT | TCAGTTGCAA | AGCTTAAGCC | TTCGCCAACA | CTTGTATCTA | ATGTTTCATC | 4380 |
| | ATCTAAAGAT | ACTGGAACAT | GTAATAATGA | AGATGATGGT | TGAATGACAA | GTTCATTAGT | 4440 |
| | GTGTGCTAAC | AATTTATCGA | TTAAGACTTT | TTTAGCTTCA | ATGTCACTTG | CCCATACATT | 450 |
| 50 | ACGACCATCA | ATAATTCCAG | CGTATAATGT | TTTTGATTTA | TCAAAATCTC | CAGCTTCAAT | 456 |
| | TTGTTTAAGG | TTATAGCCAT | TATCATGGAC | AAAGTCTAAA | CCTATACCAC | CAACAGGTAA | 462 |

| | AACACCAGCT | TTTTCGAAAT | ' AGTCATAAGC | TTCACGTGTA | ATATTTTCAT | AGCTTTCGCT | 474 |
|----|--------------|------------|--------------|--------------|------------|-------------|------|
| 5 | GTCGTCTGTA | ACTAAGATTG | GCTCATCAAC | TTGAATGTAC | TCAGCACCTG | CATCAATTAA | 480 |
| 3 | TGATTCAAAC | ACTTCTTTAT | AAAGTGGTAA | TAACGTTTTA | ACTITITCTT | CAAAAGTTTG | 4860 |
| | GTGACCGCCT | TTTGATAATT | TAACAAAAGT | ' AATCGGACCA | ACAATGACAG | GGTGAGCGTT | 4920 |
| 10 | AACGTTTAAA | GATTGGGCAT | ATTTAAAGCG | ATCTAATAAT | ACATTGCGAC | TCACTTTAGG | 4980 |
| | CTCAACATTG | TCCCATTCAG | GTACGATGTA | ATGATAGTTA | GTGTTAAACC | ATTTTATAAG | 504 |
| | TGCACTTGCA | ACATGGTCTT | TATTACCGCG | AGCAATATCA | AATAATAAAT | CATCATCAAT | 5100 |
| 15 | AGTTCTTCCT | TGGAAACGTT | CAGGGATGAT | GTTGAATAAT | AATGACGTAT | CTAATATATG | 5160 |
| | GTCATATAAA | GAGAAATCAC | CAACTGGGAT | GCTATCTAAG | TGATAGTACT | TTTGLAATAA | 5220 |
| | TAAATTTYCT | TTATGTAGAT | CAGTTAATGT | TTGATCTAAT | TCTTCTTTAG | AAATCTTCTT | 5280 |
| 20 | TGCCCAATAA | CTTTCGATGG | CTTTTTTCCA | TTCTCTTTTT | CTACCTAATC | TTGGGAATCC | 5340 |
| | TAAGTTTGAT | GTTTTAATTG | TTGTCATAAT | ATTGCCTCCT | TGTGAGCAGT | AATAGATTTT | 5400 |
| | GAGTATGCTG | CAAGTTCTAA | TGAATCTTCG | ACATTTTGAA | ACGGTGTGAT | AATGTATAAA | 5460 |
| ?5 | CCATTAAAAT | ATTCATGAAC | AGTATCGATT | AAATCCTTTG | AAAGCTTAAG | ACTTAGTTCT | 5520 |
| ٠ | CGTGTTTTGG | CTTTATCATC | TTTAACTGCT | TCAAATTGTT | GTAAAATTTC | ATCTGACATC | 5580 |
| | TTGATTCCTG | GCACTTCATT | ATGCAAAAAG | AGTGCGTTTT | TGTAACTTGC | GATAGGCATA | 5640 |
| 30 | ATGCCTATGA | AAAATGGTTT | GTTCAAGTGC | TTAGTGGCAT | GGTAAATTTC | AATGATTTTC | 5700 |
| | TCTTTGCTGT | ACACGGGTTG | TGTTATAAAA | TAAGACATTC | CGCTTTCTAT | CTTTTTCTCT | 5760 |
| 15 | AATCTTTTGA | CGGCACCATA | TAATTTACGA | ACATTAGGGT | TAAAGGCGCC | AgcGATGTTG | 5820 |
| | AAGTGTGTAC | GTTTCTTCAG | CGCATCACCG | TCAGTGTTAA | TACCTTGATT | AAATCTTAGA: | 5880 |
| | GCGĀGTTCAG | TTAATCCTTT | AGAATTAACA | TCATAGACAT | TGGTTGCACC | TGGTAAGTGA | 5940 |
| o | -CCAACTTTTG- | AAGGATCACC | AGTTATGGCT | AATATTTCGT | TAACGCCAAT | GAGCGATAAT | 6000 |
| | CCAAGTAAAT | GGGACTGCAA | GCCGATTAAG | TTTCGGTCTC | GACATGTAAT | ATGTACGAGT | 6060 |
| | GGTTCAATAT | TGTAATATTG | CTTAATTAAG | CTAGCAGCAG | CAATATTGCT | AATTCTGACA | 6120 |
| 5 | GTTGCCAATG | AATTATCTGC | GAGTGTTACC | GCATCTACAT | TAGCTTTATC | AAGTTTAGCG | 6180 |
| | ATATTTTCAA | AAAATCTATC | CGTGTCTAAA | TGTTTCGGTG | TATCCAATTC | GATAATAACG | 6240 |
| | GTTGGACGTT | CTTGAACCTT | AGATGTTAAT | GATTGTCTAA | CTTTATTTTG | AGATGGATTG | 6300 |
| 0 | AAAAGTGCTT | TCGTTGGTAT | CGGAATCACT | TTTTTGTCAT | TAACAGGTTT | AAGTGTCTGA | 6360 |
| | ATAGATTCTT | TAATAAATTT | GATGTGCTCT | GGCGTTGTAC | CACAGCAACC | ACCAATTAAA | 6420 |

| | TACTTAAATT | CACTATTTTC | AATATCTAAT | AAGCTGGCAT | TTGGATAACA | AGATAAGAAT | 6540 |
|-----|------------|------------|------------|-------------|------------|---------------------|------|
| | GCGTGCTCTG | GTAATTCAAT | ATGTGTGAAA | GACTCTTGCA | TATGGTGCGG | GCCATGATGA | 6600 |
| 5 | CAATTGAGTC | CCACGATGTT | TGCACCACAT | TGAACGAGTT | GTTTTAATCC | TTCATTGATT | 6660 |
| | GCCTGACCAT | TAACTAAGTA | ATTTGTGTTT | GAAGCGGTTA | ATTGAGCAAT | GATTGGAATG | 6720 |
| 10 | TCGTATTTCT | TTCTCGTTCG | TGAAATGACA | TTTGTTAACT | CTTCTAGGTC | GTAATACGTT | 6780 |
| ,,, | TCGAAAAGTA | GCGCGTCAAC | GCCTTCTTCA | ATTAAGGTGT | CTATTTGAAT | TTCAGTATGA | 6840 |
| | TAAAGAATAG | TTTGTAAGCT | GATATCCTCT | TGTTTGATAC | CTCTAAACCC | ACCAACTGTG | 6900 |
| 15 | CCTAATATAT | ACGTATCTTT | ATTTGCTGCT | TTTTTTGCGA | TGCGAACGGC | GGCTTGATGT | 6960 |
| | ATTGCTTTAA | CTTTATCTTC | AAGACCGAAT | CGTTTTAACT | TTTCAAAATT | TGCACCATAA | 7020 |
| | GTATTGGTTT | GAATGACATC | AGCACCGGCT | TCAATATATG | AACGATGGAT | GCGTTCAACT | 7080 |
| 20 | TTATCTGGAT | GGCTAAGATT | ATATGCTTCT | GGACAGGTGT | CTAATCCTTC | AGAGTATAAA | 7140 |
| | ATGGTTCCTA | TAGCGCCATC | AGCTACTAAA | ACATTATCTT | TCAATTGTGT | GAGGAATTGA | 7200 |
| | CTCATTGAAT | GCCTCCTTTA | ATGCGTATTT | GATGTCTGCA | ATGAGTTCAT | CAGGATCTTC | 7260 |
| 25 | GAGACCAACA | CTTAATCGGA | ATAGACCGAA | AGTGATACCA | CGTTCTTGTC | TCACTTCTTC | 7320 |
| | AGGTAGTGCA | GCGTGAGACA | TTGTTGCTGG | ATGTGAAAGG | ATCGTTTCAA | CACCGCCCAG | 7380 |
| | ACTCACTGAA | ACGAGTGGTA | ATGTCAGTGC | ATCGACAAAT | TGTTGTGCTT | TAGACTCATC | 7440 |
| 30 | AGCTAAACGA | AAGCCAATAA | CGGCACCGCC | ATTTTTAGCT | TGTTCTAAAT | GAGCAGTAGT | 7500 |
| | GAGTCCCGGA | TAATAAACTT | CTGAAATTTC | ATCTTGCTTT | attaaaaatg | ACACGATTTT | 7560 |
| ae | TTGAGCGTTT | TCGACAGATT | GTTTAAATCT | GATTGGAAAA | GTTTTTAAAT | GTTTAGCAAG | 7620 |
| 35 | TGTCCAGCTA | TCCTGAGCAG | ATAACATATT | GCCTGTACCA | TTTTGTATTA | AATAAAGAGC | 7680 |
| | GTCACTAATT | GCCTCATTAT | TAGTTATGAC | AGCACCAGCA | ATTAAATCGC | TATGTCCACT | 7740 |
| ‡O | TAAAAATTTT | GTAGCACTAT | GAATGACAAT | ATCAGCGCCA | agtaataaag | GTGATTGACc | 7800 |
| | TAACGGTGTC | ATAAATGTAT | TGTCCACAGC | TACCAGTAGT | TCATGCTTTT | CGGCTATTTT | 7860 |
| | AGAAACAGCT | TTGATATCAG | TAATTTTAAA | ACAGGGATTC | GATGGTGTTT | CGATATAAAT | 7920 |
| 15 | TAATTTTGTG | TTTGATTGAA | TGGCACCCTC | GATTTGTTCG | AGCTTTGTAG | TATCTACGGT | 7980 |
| | TGTAAATTCA | ATATTAAATC | GATTCAAAAT | TTGCTCAGTG | AGGCGAAAAG | TACCGCCATA | 8040 |
| | TACATCATCG | GGTAAGATGA | CATGATCACC | AGATTTGAAA | GTCAAAAGTA | CTGCTGAAAT | 8100 |
| 50 | AGCAGCAATA | CCTGATGCAA | AAGCAAAAGC | GAATTTTCCC | TGTTCTAATC | GTGCTAACTT | 8160 |
| | СТСТТСТВВВ | AGTTCACGCT | TACCCTTCCC | CTTCCTCC NT | 33TC3T3TT | 3 3 C 3 T C C C C 3 | 0226 |

| | TCCACACCTC | TACGCCAATC | GAATATCACT | TCTGTCTCTT | TTGAAAGTGT | CATACAATCT | 834 |
|----|------------|------------|------------|------------|------------|------------|-------|
| 5 | CTCCAATCTG | AGCTTTATCT | AATGCTTGGA | TGATATCGCG | TTCGATGTCT | TCATAATTTT | 840 |
| | CAACACCTAG | TGATAAGCGG | ATTAAATACT | CATCAATGCC | ACGTTTATCT | TTTTCAGCAT | 846 |
| | CTGGCATATC | AACATGTGTT | TGGGTGTAAG | GGAAGGTCAC | TAATGTTTCA | GTACCTCCTA | 8520 |
| 10 | AACTTTCTGC | AAAAATGCAA | ATGTCTAAAT | TTTCTAATAA | TTTAGCGACG | CTATAGGCCT | 8580 |
| | TGTTAAGTCT | TAAACTAAGC | ATGCCAGTTT | GCCCGCTATA | TAGTACTTCG | TCAATTGCTT | 8640 |
| | GAAGTGACTG | ACATTTTTTA | GCAAGTTTTC | TAGCGTTTGA | TTGCGCACGC | TCAATGCGTA | 8700 |
| 15 | AATGCAAAGT | TTTAAGTCCA | CGTAACAACA | AATAACTATC | TATTGGTGAA | AGTGTTGCGC | 8760 |
| | CAGTCATGTT | GTGAAAATCA | AACAACTGTT | GCGCGAGTGA | TTCATCTTTG | ACGGTTACGA | 8820 |
| | CACCTGCTAG | TACATCGTTA | TGTCCGCCAA | TATATTTCGT | GGCTGAATGT | AAGACTATAT | 8886 |
| 20 | CAGCACCTTC | TGCTAGTGGT | GTTGAAAGAT | AAGGTGTTAA | AAAAGTATTG | TCGATAATTG | 8940 |
| | ACAATAAGCC | TTTAGCTTTA | CAAAGTTGAT | AGTATGGCTT | TACATCAATA | GCAATCATTT | 9000 |
| | GTGGGTTAGA | TATTGGTTCA | ATGAATAATG | CAACTGTTTT | ATCAGTGATT | TCTTTTTCAA | 9060 |
| 25 | CTTGTTCATA | ATCTGTAAAA | TCAACGTACT | TAAATTTGAT | ATCGTATTGT | TGCTCGTAAA | 9120 |
| | ATTCAAATAA | TCTAAATGTG | CCACCATATA | AATCGAATGA | AACTAAAATT | TCATCATGAG | 9180 |
| 30 | GTTTAAATAG | ATTACATATT | AATTGAATGG | CTGACATTCC | ACTTGATGTA | GCGAATGATG | 9240 |
| : | CAATACCATG | CTCAAGTTTG | GCAAAACAGG | TTTCAAATGT | TGAGCGTGTA | GGATTTTTAG | 9300 |
| | TACGTGTATA | ATCAAAACCT | GTCGATTGTC | CTAGTTTTGG | ATGCTTGTAG | GCAGTAGATA | 9360 |
| 35 | AATGGATTGG | ATTCGCTATA | GCACCGGTTG | AATCATCGGT | TAATGTGATT | TGGGCTAACT | 9420 |
| | GTGTATCCTT | CATATTAAGA | CCCTCCTATA | AGAAAAATA | AAAAAAGCTT | CCGTCCTTCG | 9480 |
| | TACCCGAATG | AATCGGATAA | AAAGGACGAA | AGCTTATGTT | TCGCGGTACC | ACCTITATIT | 9540 |
| 0 | GTTATTCCAT | egetgaaata | ACCTTATTCA | GTACGCATTA | AAAGTAAATA | TGCTTACTGA | 9600 |
| | ACAATTATCA | CAATTAAAGT | CAGTAAGTAA | GGATATAGTA | ATGTGCTATC | CCATACTTAT | 9660 |
| | TAACAAAAAA | TCGTGCGTAA | AGAATCCAGT | ACGCCATTTA | ACATCAATGT | TAATACTGTA | 9720 |
| 5 | TCGCTATAAC | GGGCGAACCC | GTAGACACCT | CATATTGGCA | TCAACACTCC | AAGGCCATTT | 9780 |
| • | TCAAACACGC | TTTCAAAATC | TTCTCTCAGC | TACTAAAGAC | TCTCTGTATA | AGCAGGGTGT | 9840 |
| | GTTTTACTTY | CCTCTTTATT | GTGTTTACGT | TTCATTAAAC | TGTTATAAGA | TATTAATTAG | 9900 |
| 0 | CTTACAGAGT | AAAAAAAGAT | TTGTCAACAA | TTATTCAGAA | AATTTTGATT | TAAAAGTTAA | 9960 |
| | TTTGTTTGTG | AAATTGTAAT | TGGTATCTTG | AAGTTGAAAA | ATGAATTATT | TTTTAAATAA | 10020 |

| | TCAAATAAAA | AGTGATGTGA | GTGAATTGTC | AAAAAGTGAA | GATCAACGTA | TTACTAAAAC | 10140 |
|---|------------|------------|------------|------------|------------|------------|-------|
| | AAAAGATGAA | CAAATTAAGC | AAATAGATAT | ATCGGATATC | AAACCGAATC | CGTATCAGCC | 10200 |
| ; | CCGAAAAACT | TTCGATGAAA | ATCATTTAAA | TGATTTGGCA | GATTCAATTA | AGCAATATGG | 10260 |
| | AATTTTGCAA | CCAATTGTGC | TTAGAAAAAC | AGTTCAAGGT | TATTACATTG | TAGTTGGTGA | 10320 |
| | AAGAAGGTTT | AGAGCTTCGA | AAATTGCTGG | TCTAAAATAC | GTATCAGCGA | TTATCAAAGA | 10380 |
| 0 | TTTAACAGAT | GAAGATATGA | TGGAACTGGC | GGTCATCGAA | AATTTACAAC | GAGAAGACTT | 10440 |
| | AAATGCGATT | GAAGAAGCTG | AAAGTTATCA | ACGTTTGATG | ACAGATTTGA | AAATTACACA | 10500 |
| 5 | ACAAGAAGTA | GCGAAACGAT | TGAGTAAGTC | GCGCCCGTAT | ATAGCGAATA | TGTTGAGGTT | 10560 |
| - | ATTACATTTG | CCGAAAAAGA | TTGCTGACAT | GGTAAAAGAT | GGGCGACTGA | CAAGTGCACA | 10620 |
| | TGGACGAACG | TTATTGGCAA | TTAAAGATGA | ACAACAAATG | CTTAGGTTAG | CGAAACGGGT | 10680 |
| 0 | TGTTAAAGAA | AAGTGGAGTG | TCAGATATTT | AGAAAACCAT | GTTAATGAAT | TAAAAAATGT | 10740 |
| | TTCGTCAAAG | TCGGAAACAG | ACAAAGTAGA | TATAACTAAG | CCTAAATTTA | TAAAGCAGCA | 10800 |
| | AGAACGACAG | TTGCGAGAAC | AGTATGGTAC | CAAAGTAGAT | ATATCAATAA | AAAAATCGGT | 10860 |
| 5 | TGGTAAAATC | TCATTTGAGT | TTGATTCACA | AGAAGATTTT | GTGAGAATAA | TTGAACAATT | 10920 |
| | AAATCGTAGG | TATGGTAAAT | AGTTACACAA | TTTTATATAA | TAACTCTTTG | TGCAAGTGTA | 10980 |
| | AATAAATTGT | AATCAGTGAC | ATTTGATTCT | AGAT | | | 11014 |

(2) INFORMATION FOR SEQ ID NO: 92:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6022 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 92:

| 60 | GTTATATGTG | ATTATAGGAA | CATAATATAT | TTCTAGTTTA | GAATTTCACA | TCCCCTTATG |
|-----|------------|------------|------------|------------|------------|------------|
| 120 | TAACTTACTT | TATCGTCACA | CATTATCTAA | ACATCATAAT | AAgGTACCCT | TGTAACGCAA |
| 180 | CTTTGATATT | TGAGATTTGT | AGTGATTTGA | TATTGTTTGG | CATGGTATTA | ATGCTATAAT |
| 240 | TACATTGTTA | CTATTATTGA | ATTGACCAGA | CGGCCGATTA | TTTGTTTTGA | GATGGTACGC |
| 300 | TCGTGATTTG | CACGTCCGAT | TTTGCATCAG | TGAACTTATA | ATGATGGTCA | CAATTACAAC |
| 360 | TGGTGCTATG | TTGGCGCAAA | CACACATTAA | ATTTCATCAG | TACCATCAGT | TTGCCAGTTT |
| 42 | ATATCATCAT | ATACTGATAC | AAACCAATTC | TTCTGTTATC | AATCAAAGAT | ATTTCACAGC |

55

50

30

35

40

| | GCTGCACAAC | TTGACGCTGn | AGAACGCGAT | TTTTGAGCGT | TTAGATCCAC | ATAAGCTGGC | 540 |
|------------|------------|------------|-------------|------------|-------------|-------------|--------|
| 5 | CAGTTGTATT | GATGTTGCAA | ATATCGACAC | GCCAATCAAG | AKTATTTTAT | TAAATATAGA | 600 |
| | CCCGGCACAA | ATTACAACTA | TATTAGACGA | GCTAGATAAA | TACCATCAAG | AATTGGAAAT | 660 |
| | GATTCACCAT | TCAAATGAGT | ATAACATTGA | TATAACAGCG | CAAAATATTA | ACAAATATAC | 720 |
| 10 | TGCATTACAA | TATATATTTG | ATGCAGATGT | TAAATATATA | GCATTTGGTA | ATGACCACAA | 780 |
| | TGATATTGTC | ATGTTACAAC | ATGCTAGTAG | TGGCTATATT | ATAGGACCAT | CAGAAGCATA | 840 |
| | CACACACGCA | ATATTGAAAC | TTGATAAAAT | CAAACACATC | AATAATAATG | CACAAGCTAT | 900 |
| 15 | TTGCAAAGTC | TTAAAATCAT | ATAAATAAAA | ACACCCCTAT | CAAATGATAA | TCATTATCAA | 960 |
| | TCGATAGGGG | CTATTTTAAT | AAAATTCGTC | CTCGAACATT | TCTTCCTCTT | CATCTAATCC | 1020 |
| | AAATAATTCT | GCCATTTCTC | CATGTTCAAT | TAACATGTTT | AAATATGCAT | CGCGGAGTTC | 1080 |
| 20 | TTCTTCACTC | ATATCATTAA | TCATTTCTTT | AAGACTATCA | ATCCACATAT | TTCTGCGTAA | 1140 |
| | TTGATAGTCT | TCTTCAACTT | CGTTTAACAT | CATTATATGT | TTATTTGCTG | CTTCTGGACT | 1200 |
| | AGCTGTAAAG | AGTAATGCAA | TCATATGTTT | ACATATCACT | CGTCTTCCAT | CAGCATGAGG | . 1260 |
| 25 | ACAATTACAT | ATGGATTTTC | TAGGATGTTC | CATATCAATA | TAACAACGAT | ATACTTTGTT | 1320 |
| | GCCACTGCCC | TTTACTTCAG | CCTCATGCTG | CGTTTCTGAA | AATGATTTTA | AGTTAATGAC | 1380 |
| | GCATTCACTT | TGATAATAAT | TAAAGCCTCT | TTCTATAGAA | CGAATACTTG | CAATATCAAG | 1440 |
| 30 | TAATCCCATT | AATGaTACTC | CTTTTTATTA | TTATTTTTAA | ATAAAGAAAA | TAAAATAGAT | 1500 |
| | AAGTGTCTAG | ATTAAAATAC | TTGATTTATC | TATTTTAT | AACAAGTCTA | GAATTATCGC | 1560 |
| | ATTCTTAAAT | AACTAATATG | AAAATGCTTG | CACTAATTCt | TTTGTATAAG | GGTGTCTATC | 1620 |
| 35 | AACATTAAAT | AATTCCtCTA | TTGCAAAATC | ATCGACTATC | ATGCCATCCT | TAAGAACGAT | 1680 |
| | AATTCTATTA | ACTAAGCGTT | GTAACACGGA | TAAATCATGA | GAAATAACGA | TAAAATGATT | 1740 |
| | TAAGTTCGTA | ATCGTTTGCG | CTTTTAATAT | ATTGATTACA | TTTTGTTCAG | CTATAACATC | 1800 |
| 40 | TAAATTTGAA | GTTATCTCAT | CACATATTAA | AACGCGAGGC | TGTGCTAATA | ACGAACGCAT | 1860 |
| | GACATTAAAT | CTTTGTAATT | GTCCGCCACT | CACTTCGCTT | GGTAATTTAG | TCAATAATTG | 1920 |
| 45 | CGCGTTTAAC | TCAAAAGTAG | ATAAATGTTG | TAATAATAAT | TGATCCTGAG | CAGTATTATC | 1980 |
| | AGTTAGACCT | CTGTAATAAT | ATAACGCTTC | TTTTAATGAG | GTCTCAATCG | TCCAATCAGG | 2040 |
| | GTTAAAGCTA | GTTAAAGGGT | GTTGGAAAAT | CGGTAACACA | GCATTGTCAC | TTAAGTAAAT | 2100 |
| 5 <i>0</i> | CTCTCCTTTA | ACAGGTTTAA | ACAAGCCAAG | AACCAATGAA | GCGAGCGTAC | TTTTACCACA | 2160 |
| | GCCACTTTCG | ССТАВАВТАС | СУУСУЛЕТЕТС | TCCATCACCT | ATACTA ATAT | ጥር እጥአጥርጥጥር | 2220 |

| | CCCTCTTTAA | TIGIGITCIA | TATTTAATTA | GACGTTCAGT | ATACGGATGC | AAATGCTCAT | 2340 |
|-----|------------|------------|------------|------------|------------|------------|------|
| 5 | ACTIGAAATG | ATTAATATTA | CCTCGTTCAA | TGATTTGACC | TTCTTTTAAA | ACATAAATGT | 2400 |
| | ACTGACAATA | TTTCAATACA | TGACTTAAGT | TATGTGTGAT | AATAAATAAT | GTTTGACCAT | 2460 |
| | GTTCTAATAC | AATATGCTGT | AATAAATCCA | TCACTTGATT | ACCGTTCAAA | GCATCCAATG | 2520 |
| | ATGCAACTGG | TTCGTCTGCA | ATGATTAATT | TAGGCTCCAA | CATGAGAACG | CTTGCTATGT | 2580 |
| 10 | ATACGCGTTC | AAGTTGGCCC | CCAGAAAGTT | GGAAACTATA | TTTATTTAAT | ATATCTTTGC | 2640 |
| | TTTGTAAATT | AACCCACGAC | AAAGCCTTAT | CAACTTTGGA | CAAAGCCTCT | TCTTTACTAC | 2700 |
| | CTTTATAATG | CTTACGATAA | ATCGCAGTTA | ACTGTTTACC | TAATTTAGTA | TGGTCGTTAA | 2760 |
| 15. | AACTTTCTGC | ATAATTTTGA | GAAATATAGC | CAATTGTATG | ACCATAATAT | TGACTCAATC | 2820 |
| | TACTAACATT | TTCCCCATCA | AATTGGTACG | AATCATACGT | GCAGCTTAAA | TCAAATGGTA | 2880 |
| 20 | AATATTCAAG | TAAAGCTTTA | GCAATCAAAC | TTTTTCCAGC | GCCGCTCTCT | CCAATCAAGG | 2940 |
| 20 | CATTAATCTG | TTGACTAAAA | ATTTTCAAAT | CAATCCCTTT | AATAAGAGAT | TTCTCACTAG | 3000 |
| | TATTCTTTAT | TGTTAAATTT | TGTATATCAA | TGAGACTCAT | CATATTCACC | CCGTTGTTTC | 3060 |
| 25 | AGCAATCTAT | CTCTTAGTGC | ATCACCGGTT | AAATTAAAAA | TTAAAATAGT | TATAGCAATG | 3120 |
| | ACTGAAGCAG | GTGCAATCAA | CATAATTGGA | TGAGACGAAA | TAAAATCACG | ACCTTGTTGC | 3180 |
| | AACATAGCGC | CCCaCTCTGG | TGTTGGCGGT | TGTGCACCTA | ACCCAATAAA | TGATAGTGAA | 3240 |
| 30 | CTTATATATA | GAATGATTTT | ACCGAAATCA | ACGACCATCA | AAACGATAAT | AGCCGGTATA | 3300 |
| | ATTTTAGGTG | TTAAATGACG | TATTAATATT | GTTCTTGTTG | GTACATGAAA | TAATTGTGCC | 3360 |
| | ATTTTTATAT | AAGGCTTATT | CATTTCGCTA | TTAACTATAC | TTCTAGTCAA | CCTTGTGTAA | 3420 |
| 35 | TTCATCCATT | TTATTAATGT | AATTGAGATA | ACTAAATTCC | ATAAAGATGG | TTGAAAAAA | 3480 |
| | CTTGCTAAAG | CAATCATGAT | GATAAATTCT | GGAATACTTA | GACCAACATC | AATAAACCTT | 3540 |
| | AACACTAATC | GTTCAATCCA | CCCTTTTTTG | TATCCGGCAA | ATAGACCTAG | TGTAACACCT | 3600 |
| 40 | ATGACAACGA | TAGCTATTAA | TGTTAAAACA | GTAACAAACA | ATGTTGAACG | TGCACCGATA | 3660 |
| | ATAATTCGGG | TAAATAAATC | TCTCCCATAA | TCATCAGTTC | CTAATAAATG | CAACCAACTA | 3720 |
| | ATAGGTTCAA | AAGTTTGTGA | TAAATTGACT | TTGGTTGCAT | TTTCACTACT | GACAAAGAAT | 3780 |
| 45 | TGCAGTACAA | TTACCACAAA | AATAAATGCA | ACGAATACAA | AAAATATCAG | GTTATTCTTT | 3840 |
| | GAAAATATTT | TATGCATGAC | GGTCACTACT | TTCTGATATC | AATGGTGTAT | TGGTTTTGAT | 3900 |
| 50 | TTTTGGATTT | CCTAATTGTA | AACGCTGCTT | CGGATCAAGT | AATAACGTTA | ATAAATCAGC | 3960 |
| | AATCGTATTG | ATAATAACAA | CGAAGAAGCC | AATAAATAAC | ACGCATCCTT | GAATAACAGG | 4020 |

| | ATTTTCAATC | ACTACAGTAC | CACCTATTAG | ACTGCCAAGT | GAAATCCCTA | GTAATGGGAT | 4140 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | AATCGGCAAA | ATTGTTGGTT | TTAGTAAATC | ATGAATTAAA | ATATAACGTT | CATTCATACC | 4200 |
| 5 | GCGTAATCTT | GATGCTTGTA | CGATATTACT | TTGCAATAAC | ATCAATAAAT | TAGAACGCAC | 4260 |
| | TAAACGAATG | ATGTATGCAC | ACATACCTAA | AGATAGCGTG | ATTACAGGTA | ATATAAACTG | 4320 |
| | ACTTAGTATA | ACGCTATCTA | TATTCATTAA | ATTTGTGACA | ATAAATAATA | AAATAATACC | 4380 |
| 10 | GATAAAGAAC | GCTGGTAAAC | TAATCGATAG | TGTTGAGATC | ACTCTAATCA | CTTTATCCGT | 4440 |
| | CCACTTATGA | AATCGTTTGG | CTGCTATAAT | GCCGAGCGGT | ATAGATATGC | ATAACGACAC | 4500 |
| 15 | TACTAATGTT | GAAAATGATA | TGAGTAATGT | TATGGGTGCA | TAGTTGAATA | ATATCTGTGT | 4560 |
| , | TACCGGTTCT | TTTGATTCAA | AACTTTTTCC | ТАААТТАААА | TGTAATAAAT | GATTCATCCA | 4620 |
| | ATGCCACCAC | TGTACCAATA | AAGAATCATT | TAATCCCAAT | TTATCTTTGG | TTGCATTTAT | 4680 |
| 20 | TTGTTCCGTC | GACACTTGTG | CTACATCAAG | ATGTAATATT | TTATCAACAG | GATTGCCTGG | 4740 |
| | TGATAATTTC | ATTAAAATGA | ATGTAAGTGT | AGAAATAACA | AATAAAACAA | CTATCATTTG | 4800 |
| | CATCAGTCTA | TACAACATAG | ACTTTATTAT | GAACATAATA | GTCCCCCTCC | TTGTGTAAGT | 4860 |
| 25 | TACTAACACT | TTCTTTTTAC | ATGAGAATGG | CGCATGTATA | TGCAACTTAC | ATATTAAGAA | 4920 |
| | CTAACGTTCA | TTATAGTATT | ATCCATAAAG | AAATTGAAGT | ATATTTAATT | TTTTAACAAA | 4980 |
| | ATCATTATAA | AATATAATAT | TTTGAATCAA | GTCAACCATG | TAAAATATAA | AAAAGTCAAA | 5040 |
| 30 | ACAAAAACAA | CTATAGCACT | GTATTCCATC | TCTTTCGAAA | TAATTGTTAC | TGCAGTGTAA | 5100 |
| | CTTAAAAGTC | GATGATTTTG | TGCATATAGT | TGTCGAATAT | TATTTTTTAT | CTTTACGGCG | 5160 |
| | AAGTTCAGCG | CCCTCATAGC | CGTATTTTC | AATTTGCTTT | TCTAATTTAC | GCGCTTTTCT | 5220 |
| 35 | TTCTTTACGC | CAATTTCTAG | TAAAATACCA | TAATAGAAAA | CTAATTAATA | AACTCATAAT | 5280 |
| , | CGCTAAAAAT | GCAGCGTATC | CTAATAATGG | TTGATATTTT | ATATCTTGAA | AATTTGGAAT | 53,40 |
| | AAAAAATGCA | AGCACACCTA | ATATAACAAA | TGTAATTACT | GCAGATACAA | ACCATTTATT | 5400 |
| 40 | TAAAACTAAG | CAACAGAATA | TTGTTAATAA | AATCATTATT | AATGTTGTGA | TCCATAAATA | 5460 |
| | ATTAGGCAŤA | TCGAATAATG | TCATATTCAT | TCTCCTTTTA | TTTCATTACT | TTCCTTGTAT | 5520 |
| 45 | ACATTTTATT | TTTTTAAATA | AAAAACTTAA | ACAATAGCAG | TCAGTTTCAA | GCAATATTCT | 5580 |
| | ATCTACTAAT | ÀGAAAAATCA | TTGTTCCTTG | CGACATGGAA | ATCGTAACAT | TATCGTTTAG | 5640 |
| • | GAGACAAAAT | TATGTATAAT | GAATGTATTA | TACCAAAGGA | GTGATTATAT | GTCTCAAGGT | 5700 |
| 50 | TTACCTTTAA | GAGAAGATGT | TCCTGTTTCA | GAAACATGGG | ATTTAGTAGA | CTTATTTAAA | 5760 |
| | GATGATCAAC | AATATTATGA | AAGTATTGAC | GCTCTAGTAC | Ancaagcaaa | TCAATTTCAT | 5820 |

| | GAAAATATTT TAATTGCCTT AGATCGCTTA AGTAATTATG CAGAACTACG TTTAAGTGTA | 5940 |
|----------------|--|------|
| | GATACTAGTA ATATCGAGGC ACAAGTATTG AGCGCTAAAT TATCTACTAC ATACGGTAAA | 6000 |
| 5 | ATTGTTAAGC CAATTATCCT TT | 6022 |
| | (2) INFORMATION FOR SEQ ID NO: 93: | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 476 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 15 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 93: | |
| | CCATCAATAA TGTATACATG ATTGGCATCA TATTCCCCTT TAATTAGAGA GCTACGTACA | 60 |
| 20 | GTTTGTYTTA TTAAAGTAGA ACTAATAAAT AACCATCTCT TATGTGCACA AACACTTCCC | 120 |
| 20 | GCAACAATTG ATTCAGTTTT ACCAACCCGT GGCATACCTC TAATGCCAAT CAACTTATGA | 180 |
| | CCTTCTTCTT TGAACAATTC AGCTAAAAAG TCTACTAACA AGCCTAAATC TTCACGCTCA | 240 |
| 25 | AATCGAAAGG TTTTCTTATC TTTTGCATCT TGCTCAATAT ATCTTCCATG TCTTACTGCA | 300 |
| | AGACGGTCTC TTAATTCTGG TTTTTTAAGC TTTGTTATTT CAATTTCATT TATACCACGA | 360 |
| | GCTATTTGCT CAAAACGTTC AACTTTTTCA AGATTGTCTG TTTTAATTAA AAGGCCTCGT | 420 |
| 30 | TTACCTTGAT CAACACCATT AATTGTAACA ATACTTATAC CTAACATACC TAATAA | 476 |
| | (2) INFORMATION FOR SEQ ID NO: 94: | - |
| 35 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 3633 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 40 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94: | |
| | AGAAATACAA CGAAGCATAT AAATATAACC GATCTTTTTT CTAATTGAAT ATTAAGTAAG | 60 |
| | TGTATGTACT TTCTGGAAGT AGCACCTAGT rGGATTGTLC CTCCTACAAC AGGCCAAAAA | 120 |
| 4 5 | TTTTTATTTT TAACTGGCTT AACAGTGTTC AGTTTTTCAT ACTCTTCTCT ACTAATTTTG | 180 |
| | GCGCACCTTT TTGGAATGAA CCAATTAATA AATGGAAAAA AGTATACAAG CCAAGTTCTT | 240 |
| 50 | ATTACATCGA CCATTAAATA CTCATCATCA TACTTAATAA CTCTGTATTT CGGATTTTTA | 300 |
| - - | TTGATAATTT CGGTTTCACA AAGCAATAAT TATCACTTCC TATTAATAAC AAATTCACAC | 360 |

| | TTATATGACC | TTAAATATAT | AACATGAATC | TTTTTGTCTA | TTATTGAAGA | CATATTTATA | 480 |
|-----------|------------|------------|-------------|------------|------------|------------|------|
| 5 | AAGAAAAATA | GCATTGTCAT | AATAACCCAA | GCAATAAATA | CTATAATATT | TTGGATAGAT | 540 |
| 3 | AAACTAATCA | rracatetaa | GAACATGATT | gataatccac | CACAGAAAAA | ATAAGAAAAT | 600 |
| | AGTACAAAGC | AAAGATTCTT | GAATGATGGA | AAAATCATAA | TTTTTCCATT | GCTACTCCGA | 660 |
| 10 | TCATTATAGA | TAGATAACTT | TACTTTCTGA | ТТТАААТАТА | TATAAAACAC | TAGAATACTT | 720 |
| | AATAATAAA | CCGAACAAAT | GATAATAACG | CAATTTTTT | CTAAATGAGA | ATCAGGTATA | 780 |
| | TATATTTAT | CTCTAAACAT | AGTGCCAAAT | AAAAGTATGC | TACCTATAGC | TGGCCATAAA | 840 |
| 15 | GCTTTaTTTT | TAACTGGTTT | GACAATATTT | AAATTATCAA | AATCTTCTCT | GCTGATTTGG | 900 |
| | ACATATTTT | TTGGTATTAA | CCAATTAATA | AACGGAAAGA | ACAAAACTAA | CCAGGTGCTT | 960 |
| | ACTAAATCAA | TCATCAGATA | GTCGTTTTTA | TATTTAATAA | TTCTATATCT | GGGATTTTTG | 1020 |
| 20 | TTTACAACTC | TAACCTCGCA | AAGCAATATC | TCCACTTCCG | TCTCGTTGGT | TTTATATCTA | 1080 |
| | ATACACTTTC | AGATACTTTA | TAAGTGTTTT | GTATTTTAGT | AACATACTAT | TTTCCTGTTT | 1140 |
| 25 | ATTACTTAAC | TTACGAACTA | CAATCTAAGT | TTAGTAATTT | CTATTGCTTT | TTAAGTTTGG | 1200 |
| | CATAAACCTT | TTTATTACTA | ATTGAGCCCA | TGCTTATTAG | AAAGAAAAA | ATTGTAATAA | 1260 |
| | TAATCCACAT | AATAAATACC | AGTAGATTTT | GAGGTTTTAT | AGTCATTAGC | CATATTAAAA | 1320 |
| <i>30</i> | ATAATATAGA | ACAACCTCCT | AATAATAGAT | ATGTGAAAAC | TATAAAACTT | CCATCTTTAA | 1380 |
| | AAGTAGGCAC | TAATATAACC | CTATTTTCAT | TATCTAGATT | ATCATCATAT | ATCTTTAGTT | 1440 |
| | TAAGCTTTTT | ATTTAAGTAA | ATGTAAAATG | CTGCAATACC | TATAAATCCT | ATAAAACATA | 1500 |
| 35 | AAGATATTAA | AATCTTATTA | TCTAATTGAA | CTTCAAACGT | ATGTACATAT | TTCCGTAAAA | 1560 |
| | TAACTACAAA | TAAAAACGAA | CTACCAGTAA | CTGGCCAGAA | AATATTATTT | TTATTTTGTT | 1620 |
| 40 | TATCÁACATT | TAAATTTTCA | AGTTCCTTCT | CACTAAGTTT | TGCATACCTT | TTGGGAATGA | 1680 |
| | ACCAATTAAT | AAAAGGAAAA | AAGTATACAA | GCCAAGTGCT | TACTAAATCA | ATTAACAAAT | 1740 |
| | ACTCATCATT | ATATTGAACG | ACT/TTATATC | TCGGATTTTT | ATTAATAACC | TTAATATTAA | 1800 |
| 45 | AAAGCAAAAC | TCACCACGCC | CATTTCATTG | GATTTATATG | ATTGCTAATA | ATATTTTTAG | 1860 |
| | CTTCACTAAC | AGCATTCCCA | ACACTATCCA | TGGATTTTTC | TGTAGTTTTT | TTAACAACAT | 1920 |
| | CTATACTATT | ATCGATTTTA | TGCCCTACCC | AGTCTACTTT | ATCTTTTAAT | CCAAAAATAT | 1980 |
| 50 | TATTTTGATA | AATTAAATCT | GTTCCTAATG | CAAATACTGT | ACTCATAGCC | AAACCTGCTA | 2040 |
| | AAATCACCCA | TCCTACTGGA | TTACTTCCTA | AAACAAAAGT | CGCTAATCCA | GCTCCAACTG | 2100 |
| <i>55</i> | CTGTCCCTGC | AGATCCAGCT | GCAAGCGTgC | ATACCATTAT | GCGACAACGC | CTCTCCAAAT | 2160 |
| ~~ | | | | | | | |

| | CCTTTACCTA | GGTATTTTCC | GCCTTTTGCA | AATTTACTAC | CATTTTCTAT | AAACACATTA | 2280 |
|----|------------|------------|-------------|------------|------------|------------|------|
| | CCTGATGTAC | GTTTGACTTC | CACAAATGAA | TTTGGACCTG | CTGGGCCTTT | CACTCCACCT | 2340 |
| 5 | GCTGTATTGa | TAAATACACC | GAATTTACTT | GCATTTATAC | CGTCTTGCTC | TAAAAGTGTT | 2400 |
| | GACGTAATAT | CTAATCCTAT | ATCTCTTTTA | ATACTGTCTT | TATTGTCATT | TATATATTC | 2460 |
| o | AATATACTTT | TCGGGATATC | GTCTTCTGGA | TGTTCTTTGG | CATATGCCTT | TATAACAGCA | 2520 |
| Ü | AAGTCTGCTT | TATTTAAAGT | TTCTTTCTCT | GCTTTATGTT | CAATTTTCCC | CATAGCAACT | 2580 |
| | TTCAAATATT | TTTCATGACT | TGCTTTGGCC | CAATCAAGTT | CTTTACCTGA | AGGAATATTA | 2640 |
| 5 | AATTGATTTG | TTGAAAAGTT | CCAAAAATTC | TGCGCTTGGG | TAAGTCCTTG | TTGGACAATT | 2700 |
| | TTTTGAAATT | CTTCAACTTC | TTTAAAATATT | TCTGGTGATT | TTTGATTAAA | CTCACGCAAT | 2760 |
| | TTGCGTAGCT | TCTCTTCTAA | TTCATGTTTT | TGTTGACCTA | ATGTTCGTAT | TATTTGTTGG | 2820 |
| ro | TTCGATGAAA | TGGCTTGCTG | ATTATCGGAA | GCATGCTTTT | TCAAATTGTT | ATTCAAATTT | 2880 |
| | TCATATCGCG | TAATTTGTTG | ACTTAATGAT | CTGATATCTT | CTTCAAGCTC | TGATTCTTTT | 2940 |
| | AAAGATATGC | TATCAACCTC | ACTCGTATAA | CGTGACACAA | AATTaTCGCA | AGCTTGCTTC | 3000 |
| P5 | GTTAAATCAC | TCAATGTTTT | CATACTTGTT | GATAATGGAA | TTAACACCGT | ACTAAAAAAT | 3060 |
| | TGCTTAGCTG | ACGTATACGC | TTTCCCTTTA | AGCGCATCAT | CATTAATAAA | TTGAGTAATT | 3120 |
| 10 | GCTTTTTCCA | ACGCATCATA | ATTTGAATTC | ATTGTTTGAC | TCAAATTCCC | CACACTTGAA | 3180 |
| | GCTTGGTTTC | GAGATCTGTC | TAAATACATG | TCAATACTCA | TCGGCATGCT | CCTTTTTCAA | 3240 |
| | AAATATATGA | TTTTCAAACT | ATTTAAAATC | AAATGCTTTT | TACATCTACA | AAGTTGTAAA | 3300 |
| 35 | ATTTTAAAAC | TCGGCGATGA | TTATTTCTTA | TGTAAAGGAG | TCTAGATGCA | GGTAAATTGA | 3360 |
| | GATAACATGT | CGCCTTTTTT | CTTATTTTAG | CATATGGATA | TAATGGTGTC | TTTGTATATT | 3420 |
| | CGCAATTAAT | CAATAAAAAT | TATCTTTCAA | TATTTTAATT | TTATTGCGAC | AACATCCTTA | 3480 |
| 10 | ACATTAAATA | TATTAATATC | TCAAAATATA | TTCACTATTA | AAATATGTCA | TCAGTTGTTA | 3540 |
| | AAAGTATTTC | CTCATCATGC | GAAATATCAA | AACGTATCTA | AAATACGAAT | AAGTTTATAC | 3600 |
| | AATCACACAA | CATCATCATT | CAAAATTTTA | TTG | | | 3633 |

(2) INFORMATION FOR SEQ ID NO: 95:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2365 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

 - (D) TOPOLOGY: linear

55

45

| | TGATACGAAt | GCATTACAAT | TCATATGCAA | CATACAATTO | CTTCTACAGC | AAATGAAGTG | - 60 |
|----|------------|------------|------------|------------|------------|---------------|---------|
| | AAACAAATAG | TTGATGTGAC | ATCTGTAGCA | GAAAATGATA | CGCATTAGTC | ATAAAATTAA | 120 |
| 5 | ATGGAAATGT | CGATGAAGTG | TATCAGCAAT | TACAGCGATT | AATTAAGAAT | GCTAATGTCG | 180 |
| | AAGAGAGTGA | GAATACTGAC | AATATTAATA | GTCAAGATAC | AAGTTATACA | CCTCAAGTAA | 240 |
| 10 | AAGTAACAAC | ACCAATTTTA | GTGAAAGCAC | CAATCGCTGG | TCGTCGTATT | TTACTTAAAG | 300 |
| 10 | AAGTAAGAGA | TTCAATTTTT | AGAGAGAAAA | TGGTAGGTGA | AGGCTTAGCA | ATCAAAGCTC | 360 |
| | ATGAAGAATC | CAAAGTAATC | GCACCGTTCA | ATGGTTTAAT | ATCTATGATT | GTACCAACTA | 420 |
| 15 | AGCATGCAGT | TGGTATTCAA | TCAGAAGACG | GTGTGGACAT | AGTCATTCAT | ATTGGCGTGA | 480 |
| | ATACAGTTGA | CTTGGAAGGT | AAAGGGTTCA | AGTGCTTTGT | AAAGCAAAAT | GATCATGTTG | 540 |
| | AAGCAGGGCA | AACGTTGTTG | CAATTCGACC | AGCAATATAT | ACAACAACAA | GGCTACAATG | 600 |
| 20 | CTGACGTTAT | TGTCGTTATT | AGCAACTCTG | CCGATTTAGG | AAAAGTAGAA | CTGACAATGA | 660 |
| | ATGAAATCAT | TACGACTGAA | GATGTTATTT | TTAAAATATT | TAAAAACTAG | GAGTGTGTTG | 720 |
| | TAATAATGAC | AAAATTACCG | CAAAATTTCA | TGTGGGGTGG | CGCTCTTGCC | GCAAATCAAT | 780 |
| ?5 | TTGAAGGTGG | ATATGATAAA | GGTGGTAAAG | GGTTAAGTGT | AATTGATGTT | ATGACGAGTG | 840 |
| | GTGCACATGG | CAAAGCACGT | CAGATTACAG | AATCTATAGA | TCCCAATCAC | TATTATCCAA | 900 |
| 10 | ATCATGAAGG | TATTGATTTT | TATCATCGTT | ATAAGGAAGA | TATTGCCTTG | TTTAAAGAAA | 960 |
| | TGGGATTGAA | ATGTTTACGT | ACGTCGATTG | CGTGGACACG | TATCTTTCCG | AATGGGGATG | 1020 |
| | AAGATGTGCC | AAATGAAGAA | GGACTCGCCT | TTTATGATCG | TATCTTTGAT | GAATTAATTG | 1080 |
| 5 | CACAAGGTAT | TGAACCTGTT | GTGACGTTAT | CACATTTTGA | GATGCCACTT | CATTTAGCGA | 1140 |
| | AACATTATGG | TGGATTTAGA | AATAGAGAAG | TTGTCGATTA | TTTTGTGCAT | TTTGCGCGTG | 1200 |
| | TTGTATTTGA | AAGATATAAA | GATAAAGTTA | CATATTGGAT | GACGTTTAAT | GAAATTAATA | 1260 |
| 0 | ATCAGATGGA | CACATCAAAT | CCTATCTTTT | TATGGACGAA | TTCTGGGGTA | GCATTGACAG | 1320 |
| • | AAAATGATAA | TCCTGAAGAA | GTCYTGTATC | AAGTAGCACA | TCATGAACTT | TTAGCCAGTG | 1380 |
| | CYTTAGCAGT | TCGTCTTGGT | AAAGaGATtA | ATCCGAaGTT | TAAGATTGGr | ACmATGATTt | 1440 |
| 5 | CAMATGTACC | CmTTTATCCa | TAWTCGTGTC | ATCCGAAAGA | TATGATGGAA | GCACAAATTG | 1500 |
| | CGAATCGCTT | ACGTTTCTTT | TTCCCGGATG | TCCAAGTGAG | AGGTTATTAT | CCAAGCTATG | 1560 |
| o | СТАААААААТ | GTTGGCACGA | AAAGGATATG | ATGTTGGATG | GCAAGAAGGG | GACGACAGTA | 1620 |
| - | TTTTACAGCA | GGGCACGGTT | GATTATATTG | GCTTTAGTTA | TTACATGTCT | ACGGCTGTAA | 1680 |
| | AACATGATGT | TGATACTACA | GTTGAAAACA | ACATCGTCAA | СССТССТТТС | እ እጥር እጥጥር ጥር | 1740 |

| GATATACATT | GAATGTGTTA | TATGATCGTT | ATCAGTTACC | ACTTTTTATT | GTGGAAAATG | 1860 |
|------------|------------|------------|------------|------------|------------|------|
| GTTTTGGTGC | AGTTGATGAA | GTGGTAGATG | GACATATTCa | TGATGATTAT | CGCATTGAAT | 1920 |
| ATTTAAAAGC | ACATATTACA | GCAGCGATAG | AAGCAGTTGA | TCAAGATGGT | GTAGATTTAA | 1980 |
| TCGGTTATAC | ACCGTGGGGA | ATCATTGATA | TTGTTTCATT | TACAACCGGT | GAAATGAAGA | 2040 |
| AACGCTATGG | TTTAATATAT | GTTGATCGAG | ATAATGATGG | TCATGGCACG | ATGGAACGCT | 2100 |
| TGAAAAAAGA | TTCGTTCTAT | TGGTATCAAC | AAGTGATAGC | ATCAAATĠGA | GATAAATTAT | 2160 |
| AAAGGTATAT | TATAAGTATT | TTAGGGTTAG | AGCCCGAGAC | ATAAATTAAT | ATAGTAGGAC | 2220 |
| CTACAGTGTT | ATAATGGCGG | gccccaaca | CAAAGAATTT | CGAAAAGAAA | TTCtAcAGGT | 2280 |
| aATGCaAGtT | GCGGGGCCC | AACACAGAGA | AATTCGAAAA | GAAATTCTAC | AGGTAATGCA | 2340 |
| AGTTGGGGAA | GGACAGAAAT | TTAAA | | | | 2365 |
| | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 96:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11050 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:

| 60 | AATTTAGAGC | TATCATTACC | ACAAAAAAGT | AGTGGGGAAA | ATTTGTTGAA | CTGCGATACG |
|-----|------------|------------|------------|------------|------------|------------|
| 120 | CACTTTATCA | TCACATGTAG | GTACACACAT | GGTAATAAAG | AGCTTTGATT | AGGCATACGA |
| 180 | TCTGCTGACA | GAAAAAGTTG | TGATTCAAAT | GTTTCTCCGT | CATTAAATAT | CGCGACAAAA |
| 240 | GTAGAAATAA | TTTATTTTGG | GACAGAAAAC | TTTAACTTAT | TTTGAAGGAG | CTTTGCAAGG |
| 300 | ATAGTATATT | AATATACCAT | AGATTACATT | AGGCATTAGC | CATGAACTAA | TAAATTAGAA |
| 360 | TTATTTATTT | AAAGGCCAAG | ACATTATACG | GTTTTGTCAG | CAAAGTGAAT | ACAACCATAT |
| 420 | TTAGAGAACA | GGTAACATTA | TTTAATTGAA | ATATTTACTT | GAAAGTAGCA | TTCGCCACAA |
| 480 | TATTTCCAAT | GAGCAAGTAT | TTTTAATAAA | TATATCGTTA | AATGGAGATG | TTACAATCAA |
| 540 | CCGATTGTAC | ACAGCATTAA | CGAATTGTGT | AAGAGGTTAA | TTTCATCCGA | CAGTAACTTA |
| 600 | ATGATATATT | AAAGCTAATG | CTTTTTGTGC | AATTGATGGC | TTGCCTAGAG | AGTTCTTGGA |
| 660 | ATAACATGGC | CACATGAACT | TGAGCAGCAA | TAAATGATAA | TTTGCATTAA | TTTGACACTT |
| 720 | TATGTCAGAC | ATATGCCATC | TATCAAATTG | AAGATCGAAT | AAATTTGCTA | ATTAACAAGT |
| 780 | TTCAACtCAT | TTTTTAACTA | AATCAAACAG | AATTTTATGA | GATCAAGATG | AGTAGGATAC |

55

50

5

10

15

20

25

30

35

40

| | TGAAAAACTT GTTGTTAAAG ATCATAAAAA TTGGTTAGTA AGCAAACATT TATTCAATG | A 90 |
|----|---|--------|
| | TGTATGTGTT TAATATACAA TGTAAAATGA ATAAGTTGAA CATGAGGTCT AACGTACAT | Т 96 |
| 5 | TATACGTTAG GCCTTTTTTG CTAGCATGAT GAATAATTTA AAATGTTAGT TAAATTTGA | T 102 |
| | TGTTGAAATT ACAGTAAAAT TTAAGGTGAT GAAAAATTTA GAACTTCTAA GTTTTTGAA | A 108 |
| | AGTAAAAAT TTGTAATAGT GTAAAAATAG TATATTGATT TTTGCTAGTT AACAGAAAA | T 1140 |
| 10 | TTTAAGTTAT ATAAATAGGA AGAAAACAAA TTTTACGTAA TTTTTTTCGA AAAGCAATT | G 1200 |
| | ATATAATTCT TATTTCATTA TACAATTTAG ACTAATCTAG AAATTGAAAT GGAGTAATA | T 1260 |
| 5 | TTTTGAAAAA AAGAATTGAT TATTTGTCGA ATAAGCAGAA TAAGTATTCG ATTAGACGT | Г 1320 |
| | TTACAGTAGG TACCACATCA GTAATAGTAG GGGCAACTAT ACTATTTGGG ATAGGCAATG | 2 1380 |
| | ATCAAGCACA AGCTTCAGAA CAATCGAACG ATACAACGCA ATCTTCGAAA AATAATGCAA | A 1440 |
| 0 | GTGCAGATTC CGAAAAAAC AATATGATAG AAACACCTCA ATTAAATACA ACGGCTAATC | 3 1500 |
| | ATACATCTGA TATTAGTGCA AACACAAACA GTGCGAATGT AGATAGCACA ACAAAACCAA | 1560 |
| | TGTCTACACA AACGAGCAAT ACCACTACAA CAGAGCCAGC TTCAACAAAT GAAACACCTO | 1620 |
| 5 | AACCGACGGC AATTAAAAAT CAAGCAACTG CTGCAAAAAT GCAAGATCAA ACTGTTCCTC | 1680 |
| | AAGAAGCAAA TTCTCAAGTA GATAATAAAA CAACGAATGA TGCTAATAGC ATAGCAACAA | 1740 |
| | ACAGTGAGCT TAAAAATTCT CAAACATTAG ATTTACCACA ATCATCACCA CAAACGATTT | 1800 |
| 0 | CCAATGCGCA AGGAACTAGT AAACCAAGTG TTAGAACGAG AGCTGTACGT AGTTTAGCTG | 1860 |
| | TTGCTGAACC GGTAGTAAAT GCTGCTGATG CTAAAGGTAC AAATGTAAAT GATAAAGTTA | 1920 |
| 5 | CGGCAAGTAA TTTCAAGTTA GAAAAGACTA CATTTGACCC TAATCAAAGT GGTAACACAT | 1980 |
| _ | TTATGGCGGC AAATTTTACA GTGACAGATA AAGTGAAATC AGGGGATTAT TTTACAGCGA | 2040 |
| | AGTTACCAGA TAGTTTAACT GGTAATGGAG ACGTGGATTA TTCTAATTCA AATAATACGA | 2100 |
| o | TGCCAATTGC AGACATTAAA AGTACGAATG GCGATGTTGT AGCTAAAGCA ACATATGATA | 2160 |
| | TCTTGACTAA GACGTATACA TTTGTCTTTA CAGATTATGT AAATAATAAA GAAAATATTA | . 2220 |
| | ACGGACAATT TTCATTACCT TTATTTACAG ACCGAGCAAA GGCACCTAAA TCAGGAACAT | 2280 |
| 5 | ATGATGCGAA TATTAATATT GCGGATGAAA TGTTTAATAA TAAAATTACT TATAACTATA | 2340 |
| | GTTCGCCAAT TGCAGGAATT GATAAACCAA ATGGCGCGAA CATTTCTTCT CAAATTATTG | 2400 |
| | GTGTAGATAC AGCTTCAGGT CAAAACACAT ACAAGCAAAC AGTATTTGTT AACCCTAAGC | |
|) | AACGAGTTTT AGGTAATACG TGGGTGTATA TTAAAGGCTA CCAAGATAAA ATCGAAGAAA | |
| | GTAGCGGTAA AGTAAGTGCT ACAGATACAA AACTGAGAAT TTTTGAAGTG AATGATACAT | |

| | ACCAATTTAA | AAATAGAATC | TATTATGAGO | ATCCAAATGT | AGCTAGTATT | AAATTTGGTG | 270 |
|----|------------|------------|------------|------------|-------------------|------------|------|
| | ATATTACTAA | AACATATGTA | GTATTAGTAG | AAGGGCATTA | CGACAATACA | GGTAAGAACT | 276 |
| 5 | TAAAAACTCA | GGTTATTCAA | GAAAATGTTG | ATCCTGTAAC | AAATAGAGAC | TACAGTATTT | 282 |
| | TCGGTTGGAA | TAATGAGAAT | GTTGTACGTT | ATGGTGGTGG | AAGTGCTGAT | GGTGATTCAG | 288 |
| | CAGTAAATCC | GAAAGACCCA | ACTCCAGGGC | CGCCGGTTGA | CCCAGAACCA | AGTCCAGACC | 294 |
| 10 | CAGAACCAGA | ACCAACGCCA | GATCCAGAAC | CAAGTCCAGA | CCCAGAACCG | GAACCAAGCC | 3000 |
| | CAGACCCGGA | TCCGGATTCG | GATTCAGACA | GTGACTCAGG | CTCAGACAGC | GACTCAGGTT | 3060 |
| 15 | CAGATAGCGA | CTCAGAATCA | GATAGCGATT | CGGATTCAGA | CAGTGATTCA | GATTCAGACA | 3120 |
| | GCGACTCAGA | ATCAGATAGC | GACTCAGAAT | CAGATAGTGA | GTCAGATTCA | GACAGTGACT | 3180 |
| | CGGACTCAGA | CAGTGATTCA | GACTCAGATA | GCGATTCAGA | CTCAGATAGC | GATTCAGACT | 3240 |
| 20 | CAGACAGCGA | TTCAGATTCA | GACAGCGACT | CAGATTCAGA | CAGCGACTCA | GACTCAGATA | 3300 |
| | GCGACTCAGA | CTCAGACAGC | GACTCAGATT | CAGATAGCGA | TTCAGACTCA | GACAGCGACT | 3360 |
| | CAGACTCAGA | CAGCGACTCA | GACTCAGATA | GCGACTCAGA | TTCAGATAGC | GATTCAGACT | 3420 |
| 25 | CAGACAGCGA | CTCAGATTCA | GATAGCGATT | CGGACTCAGA | CAGCGATTCA | GATTCAGACA | 3480 |
| | GCGACTCAGA | CTCGGATAGC | GATTCAGATT | CAGATAGCGA | TTCGGATTCA | GACAGTGATT | 3540 |
| 30 | CAGATTCAGA | CAGCGACTCA | GACTCGGATA | GCGACTCAGA | CTCAGACAGC | GATTCAGACT | 3600 |
| 30 | CAGATAGCGA | CTCAGACTCG | GATAGCGACT | CGGATTCAGA | TAGCGACTCA | GACTCAGATA | 3660 |
| | GTGACTCCGA | TTCAAGAGTT | ACACCACCAA | ATAATGAACA | GAAAGCACCA | TCAAATCCTA | 3720 |
| 35 | AAGGTGAAGT | AAACCATTCT | AATAAGGTAT | CAAAACAACA | CAAAACTGAT | GCTTTACCAG | 3780 |
| | AAACAGGAGA | TAAGAGCGAA | AACACAAATG | CAACTTTATT | TGGTGCAATG | ATGGCATTAT | 3840 |
| | TAGGATCATT | ACTATIGTTT | AGAAAACGCA | AGCAAGATCA | TAAAGAAAAA | GCGTAAATAC | 3900 |
| 40 | TTTTTTAGGC | CGAATACATT | TGTATTCGGT | TTTTTTGTTG | AAAATGATTT | TAAAGTGAAT | 3960 |
| | TGATTAAGCG | TAAAATGTTG | ATAAAGTAGA | ATTAGAAAGG | GGTCATGACG | TATGGCTTAT | 4020 |
| | ATTTCATTAA | ACTATCATTC | ACCAACAATT | GGTATGCATC | AAAATTTGAC | AGTCATTTTA | 4080 |
| 45 | CCGGAAGATC | AAAGCTTCTT | TAATAGCGAT | ACAACTGTTA | AACCATTAAA | AACTTTAATG | 4140 |
| | TTGTTACATG | GATTATCAAG | TGATGAAACG | ACATATATGA | GATATACAAG | CATAGAAAGG | 4200 |
| 50 | TATGCGAATG | AACACAAATT | AGCTGTGATT | ATGCCCAATG | TGGATCATAG | CGCATATGCT | 4260 |
| 50 | AACATGGCAT | ATGGTCATAG | CTATTATGAT | TATATTTTGG | aagtgtatga | TTATGTTCAT | 4320 |
| | CAAATATTTC | CACTTTCCAA | AAAGCGTGAT | GACAATTTA | ТАССАССТСА | СТСТАТСССА | 4380 |

| | | | | | | TTTTTCAAAA | |
|----|------------|--------------|-------------|--------------|------------|--------------|------|
| | | | | * | | GTATTACTTG | |
| 5 | | | | | | GTGTGGTAAA | |
| | CAAGACTTT | TATATCAAGA | A CAACTTAGA | TTTATCGATT | ATTTATCACG | CATAAATGTT | 468 |
| 10 | CCTTATCAA1 | TTGAAGATGG | ACCAGGAGA | r catgattatg | CATATTGGGA | TCAAGCGATT | 474 |
| | AAGCGTGCTA | TAACATGGAT | GGTGAATGA: | TAATTATTTC | TTGGAAAATA | TGTGGCTGCA | 480 |
| | TTAAATACAC | AGAGTGAGAG | ATACAAACT | TTTACGCACG | ACTAACATTT | CTAAGTGTTT | 486 |
| 15 | AAATTATTTT | TGTATTAATA | TGATTGGCG | AATTTGCTGA | TACACAAAAA | TGTTTCTCGT | 492 |
| | GAAACTTAGA | TTTAGCTTAT | AGTTTTATCA | TCATTTGTAT | GACTTACATT | ATAAATTTTA | 498 |
| | TTATAATGAG | GTTAACGCTT | TGAAAGGAGI | CATCATCATG | TCGACCAATA | AAAACGATTA | 504(|
| 20 | TGAGCATATG | TTGTTTTATT | TTGCATATAA | AACCTTTATT | ACTACCGCTG | ATGAAATTAT | 510 |
| | AGAGAAGTAT | GGTATGAGTC | GTCAGCATCA | TCGTTTTTTG | TTTTTTATCA | ATAAATTACC | 5160 |
| | TGGTATTACT | ATTAAATCAT | TACTAGAAAT | ATTAGAAATT | TCTAAmCAAG | GATCACATGC | 5220 |
| 25 | AACACTTCAA | AAATTAAAAG | AGCAAGGTCT | CATTATTGAA | AAAGTTTTAG | AGACTGATCG | 5280 |
| | ACGTGTCAAA | AAATTATATT | CGACGGATAA | AGGCGATCAA | CTCATTGCTG | AATTGAACAA | 5340 |
| 30 | GGCGCAAGAT | GAATTATTGC | AAAATATATA | TCAACAAGTC | GGTTCGGATT | GGTATGATGT | 5400 |
| 30 | GATGGAAGCA | TTGGCTAAAG | GgCGACCTGG | CTTTGATTTT | ATTAAGCATT | TGAAAGATGA | 5460 |
| | AAAAGAAAGC | TAGCATCAGA | AATGTTAAAA | ATCTTCGCAT | TCTTAAATTT | AAAAAATATG | 5520 |
| 35 | TCAAAAAGTG | TATAATAAAA | ACATATAATT | TAATTGAACT | CAGTTTCAAC | ACATCTTAGA | 5580 |
| | AAGGAGTTTG | AATGATGAAA | AAATTAGCAG | TTATTTTAAC | ATTAGTTGGC | GGTTTATACT | 5640 |
| | TCGCATTTAA | AAAATACCAA | GAACGTGTTA | ACCAAGCACC | TAACATTGAG | TACTAAATTA | 5700 |
| 40 | AACCATAAAA | -AATTCCCGAA- | CACCTTGTTA | TAGTGCTCGG | GAATTTTTTT | ATGCTTTACT | 5760 |
| | TGAATATATC | AAATATTATT | TTTGCGCTTT | CTGTATTTTC | GATATTACCA | CTAAATGATT | 5820 |
| | CTGATCTAGG | TCCGTAAGCG | TAgGTATTAA | CATCCTCGCC | TGTATGTCCA | TCGGAAGTCC | 5880 |
| 45 | ACCCTGTATA | AGATTTATCA | TTTACTGGCT | TCTGAATAGC | GTGTTGTAGG | GCTTTTGTTT | 5940 |
| | GCGTTTCTAC | TTCTGCGGAT | TTTTCGTCTT | TTTCTTTTTT | AAGTAGTCTT | TTTAGCTTTT | 6000 |
| 50 | TATTCTCTTT | TTTAACCTTT | TTCATATCAT | CTTGTGAAAA | TTCAAATCCA | TAACCTTCAT | 6060 |
| 50 | TAATAACTTT | TTCAGGGTCT | TCACCTTTAG | CCATTTTTTC | TGTCATATAT | GATCCAGAGT | 6120 |
| | GTTTCATAGA | TTTAATCGGT | TGAGGATTCC | ATTCGTATCC | TTTATCTTTA | ССА А ТТСТТА | 6190 |

| | ATTGAATGGC | GTCATCGAAT | GCTTTTTCAA | AACCTTCCAT | TTCAGACATA | ACGCCTGTAA | 6300 |
|------------|------------|------------|------------|------------|-------------|------------|-------|
| | TATCGTTGGA | ATGCGCTGAT | TTATCTATAG | AAGCACCTTC | GACCATTAAA | AAGAATCCTT | 6360 |
| <i>5</i> | TTTTATTGCG | CTCAAGCTTA | CTAAGTGCAC | TTTGTTGCAT | ATCAGCTAAT | GATGGTTCGT | 6420 |
| | CTTTAGAAGC | ATCTATTGCA | AGTGGCATAT | TTTTATCTGC | AAACAAACCA | AGAACTTTAT | 6480 |
| | CTTTATCAGA | TTTTGATAAC | TCCTTACTGT | TCGTGGCAAG | GTCGTAACCA | TCTTTTTGA | 6540 |
| 10 | ATTTTTTATC | TAAATTGCCA | TTACTTTTAC | CGAAATATTT | AGCGCCGCCG | CCTAATAAAA | 6600 |
| | CATCAACTTT | ATGCTTTCCG | TTGATTTTAT | CTTTATAAAA | TTGTTTAGCG | ATTTCGTTTT | 6660 |
| 15 | TATCATCTCT | AGAAGTCACG | TGTGCAGCAT | ATGCTGCTGG | TGTTGCATCT | GTTAATTCAG | 6720 |
| | CTGTTGAAAC | AAGACCAGTC | GACTTACCTT | TTTCTTTTGC | ACGTTCAAGC | ACCGTCTTTA | 6780 |
| | CTTTCTGCTT | GTTACTGTCA | ACACCGATGG | CACCATTATA | TGTCTTATGA | CCAGAACTAA | 6840 |
| 20 | AGGCTGTTCC | GCCAGCTGCA | GAATCAGTAA | TATTCTGTTT | TGGGTCATTT | GAATATGTAC | 6900 |
| | GATTTGTGCC | TTTTAAATAT | GAATCAAAAG | CAGTAGGGGT | CATTTCTTTA | GCATGCGGAT | 6960 |
| | CATTTTTATA | ATAACGATAA | GCTGTGTTAA | ATGATGGACC | CATGCCATCG | CCAACTAÁAA | 7020 |
| 25 | AGATAACATT | TTTTGGATTT | TTAGTATTAC | CAACCGCGAA | ACTTTCATCT | TTAGAACTTT | 7080 |
| | TATCGGATTG | CGCAATTGCA | GGTGTGACAG | AACTAAAAAC | CGTTGACACG | ATAATAAGGT | 714 |
| 20 | TAGCAACTGC | AAATTTTGTG | GCTTTTTTAA | CTGATAACAT | AAGACATCCT | CCTGAGTATA | 7200 |
| 30 | TGACTATGTC | TTCAGTGTAA | AAGAGGAATT | TEGAGCAATT | ATGTAGTTTT | AGTTANAAAT | 726(0 |
| | ATGTAAACAG | AGTGATTTAG | AATAACAAAA | aATGAATATA | TATGACAATT | TGTTATAGAA | 7320 |
| 35 | AGCGTTÁGAA | TAGAAGCGTG | TGAAAATATA | GAATTAAATA | TAATTTGAGG | TGGAAAAATG | 73,80 |
| | ATACTAGTAA | TGTTATCTCC | ATTATTAATC | ATATTCTTTA | TAGTGTTGTC | TATTTTAGAA | 7440 |
| | GAGCGTAAAC | GTACGAAGAA | AAAGCAACTC | GAGAAAGAAA | AAGCAAATAC | ACTAAATCAA | 7500 |
| 40 | AATACAAATG | ACACGGAAAG | TTCAAATCAA | GAGCCGTCAT | TGCAGCAGGA | TAAAGAACAA | 7560 |
| | AAAGATAACA | AAGGATAATT | CAATTGAAGG | AAGAAGATTA | TAGATGAAAA | TATTAATTGT | 7620 |
| | TGAAGATGAT | TTTGTTATAG | CAGAGAGTTT | AGCATCTGAA | CTTAAAAAAAT | GGAATTACGG | 7680 |
| 45 | TGTTATTGTC | GTTGAACAAT | TTGATGATAT | ACTGTCTATC | TTTAACCAAA | ATCAACCTCA | 7740 |
| | GCTTGTATTG | CTAGATATTA | ATTTGCCAAC | GTTAAATGGT | TTTCATTGGT | GTCAAGAAAT | 7800 |
| 5 <i>0</i> | CCGAAAAACA | TCTAATGTGC | CAATTATATT | TATTAGTTCC | CGTATTGATA | ATATGGACCA | 7860 |
| | AATTATGGCA | ATACAAATGG | GGGGAGATGA | TTTTATCGAA | AAGCCATTTA | ACTTGTCATT | 7920 |
| | AACGATTGCC | AAAATTCAAG | CATTATTGAG | ACGAACTTAT | GACTTGTCAG | TAGCTAATGA | 7980 |

| | ACAAAACATA | CAGCTATCTT | TGACTGAATT | ACAAATATTA | AAGTTATTAT | TTCAAAATGA | 8100 |
|-----------|------------|------------|------------|------------|------------|------------|------|
| | AGaTAAATAT | GTAAGTAGrA | CTGCTTTAAT | TGAAAAATGT | TGGGaATCAG | AAAACtTCAT | 8160 |
| 5 | AGATGATAAC | ACATTAGCTG | TTAACATGAC | GCGCCTGCTG | AAAAAATTAA | ATACTATTGG | 8220 |
| | CGTTAATGAT | TTTATCATTA | CAAAGAAAAA | TGTCGGATAT | AAAGTATAGG | GTGAATGCAA | 8280 |
| | TGACCTTTCT | TAAAAGTATT | ACTCAGGAAA | TAGCAATAGT | CATAGTTATT | TTTGCTTTGT | 8340 |
| 10 | TTGGCTTAAT | GTTTTACCTG | TATCATTTGC | CATTAGAAGC | ATATTTACTA | GCACTTGGCG | 8400 |
| | TTATTTTATT | ATTATTACTC | ATATTCATAG | GTATTAAATA | TTTAAGTTTT | GTAAAAACTA | 8460 |
| 15 | TAAGCCAACA | ACAACAAATT | GAAAACTTAG | AAAATGCGTT | GTATCAGCTT | AAAAATGAAC | 8520 |
| | AAATTGAATA | TAAAAATGAT | GTAGAGAGCT | ACTTTTTAAC | ATGGGTACAT | CAAATGAAAA | 8580 |
| | CACCCATTAC | TGCAGCACAA | CTGTTACTTG | AAAGAGATGA | GCCTAATGTT | GTTAATCGTG | 8640 |
| 20 | TTCGTCAAGA | GGTTATTCAA | ATTGaTAACT | ATACAAGTTT | AGCACTTAGT | TATTTAAAGT | 8700 |
| • | TATTAAATGA | AACTTCTGaT | ATTTCTGTCA | CTAAAATTTC | GATTAATAAT | ATCATTCGCC | 8760 |
| | CAATTATTAT | GAAATATTCA | ATACAGTTTA | TTGATCAAAA | AACAAAAATC | CATTATGAAC | 8820 |
| 25 | CTTGTCATCA | CGAAGTATTA | ACTGACGTTA | GATGGACCTC | TTTAATGATA | GAACAATTAA | 8880 |
| | TAAATAATGC | ACTTAAGTAT | GCGAGAGGTA | AAGATATATG | GATTGAATTT | GATGAGCAAT | 8940 |
| 30 | CCAATCAATT | ACACGTAAAA | GATAATGGTA | TCGGTATTAG | TGAAGCGrAC | TTGCCTAAAA | 9000 |
| 30 | TATTTGATAA | GGGCTATTCA | GGTTATAATG | GCCAGCGCCA | AAGTAACTCA | AGTGGGaTTG | 9060 |
| | GTTTATTTAT | CGTAAAACAA | ATTTCAACAC | ACACAAACCA | TCCTGTTTCA | GTCGTATCTA | 9120 |
| 35 | AACAAAATGA | GGGTACAACA | TTTACGATTC | AATTTCCAGA | TGAATAAAAA | CTTTCAATAT | 9180 |
| | TGTAAGTATA | CTAGTAACAT | TTTTTTACTA | ATTTAAATGT | TATTAGTATT | TTTTTGTTTT | 9240 |
| | AATATAGAAC | TAACAAAGAA | ATGAGGTGCA | TGCCATGTTG | CTAGAAGTGn | AACATGTAAA | 9300 |
| 10' | AAAGGTTTAT | GGTAAAGGTT | TGAATGCTAC | GACAGCACTT | AATCAAATGA | ATTTATCAGT | 9360 |
| | TGGAGCTGGT | Gaatttgttg | Caattatggg | TGAGTCTGGG | tCAGGGAAGT | CTACACTACT | 9420 |
| | AAATTTAATT | GCLTCTTTTG | ATGGACTAAC | TGAAGGTGAC | ATTATTGTGG | ATGGCGCACA | 9480 |
| 15 | TTTAAATAAT | ATGAAAAATA | AAAGTAAAGC | ATTGTATCGT | CaACAAATGG | TAGGTTTTGT | 9540 |
| | TTTTCAAGAT | TTTAATCTTT | TACCAACAAT | GACGAATAAA | GAAAATATAA | TGATGCCATT | 9600 |
| 50 | AATTTTAGCT | GGTGCTAAAC | GAAAAGATAT | AGAACAAAGG | GTACATCAGT | TGGCAGTACA | 9660 |
| - | ATTACATTTA | GAGGGATTCT | TAAACAAGTA | TCCTTCTGAA | ATCTCTGGGG | GTCAGAAGCA | 9720 |
| | ACGCATTGCC | ATTGCACGTG | CATTAGTTAC | TAAGCCGACG | ATTITACTAG | CCGATGAACC | 9780 |

| | TCAATTGGAA | CAGACAATTT | TAATGGTAAC | TCATTCAAAT | ATCGATGCGT | CTTATGCAGA | 9900 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | GCGAGTCATT | TTTATTAAAG | ATGGGCGTCT | ATATCATGAA | ATATATCGTG | GTGAAGAAAG | 9960 |
| 5 | TCAATTAGCT | TTTCAACAAC | GAATAACAGA | TAGCTTAGCA | CTTGTGAATG | GAGGAAGTGT | 10020 |
| | CAATATATGA | AGTTAAGATT | GTTATGNACA | TAGTGCGACG | TCAATTTATT | ACGCAGCGAC | 10080 |
| | TTGTAATCAT | TCCATTCATT | TTAGCGGTAA | GTGTACTATT | CATGATTGAA | TATACGCTTG | 10140 |
| 10 | TGTCAATTGG | GTTAAATAGC | TACATAAAAC | AGAAGAATGA | CTTCCTAGTA | CCATTTATTA | 10200 |
| | TCATAGCTAA | TTTTTTTATG | GCGCTTTTAA | CTTTTATTTT | TATTTTCTAT | GCAAATCACT | 10260 |
| 15 | TTATGATGTC | ACAAAGACGA | AAAGAGTTTA | GCATTTTTAT | GACATTGGGC | ATGACCAAGA | 10320 |
| | AAAGTATGCG | TTTAATTGTA | GTGATGGAAA | CTATCTTACA | ATTTGTGATA | ATTTCAGTCG | 10380 |
| | TTAGTATTGC | CGGCGGATAC | TTACTTGGTG | CGATATTTTT | CTTGTTTATA | CAGAAAATAA | 10440 |
| 20 | TGGGCAGTGA | AGTTGCGACG | TTAAGGTATT | ATCCATTTGA | CTCTGTAGCG | ATGTTTATTA | 10500 |
| | CTTTGATTAT | CATTGCTGTA | TTAATGGGCA | TGCTACTTAT | ATTCAACTTG | TTTAGTATTA | 10560 |
| | ATTTTCAACG | GCCGATAACT | TATCAACATC | GTTCCGATTC | TAGTGTCATA | TCACGATGGT | 10620 |
| 25 | TGCGTTACGT | TTTAATTGTT | ATAGGAAGCG | CANACTATAT | TTAGGTTACT | TTATTGCATT | 10680 |
| | ACAACAAGAT | ACGACGTTTG | GTGCCTTTTT | TAAAATATGG | ATTGTCATAG | GATTAGTTAT | 10740 |
| | TATCGGTACT | TATGCATTTT | TTGTAGGTAT | AAGTGAAATA | ATTATTAGTA | TATTGCAGCA | 10800 |
| 30 | GGTATCAAAA | GTTTACTATC | ATCCACGGTA | TTTTTTTGTG | GTAGTTGGGA | TGCGTGTACG | 10860 |
| | TCTTAAAATG | AATGCAGTCA | GTCTTGCAAC | AATCACTTTG | CTGTGTACAT | TTTTGATTGT | 10920 |
| 35 | AACGCTCACA | ATGACATTAA | CAACCTATCG | TGATATGAAT | CATACCATTA | CGAAATTGAT | 10980 |
| | TACGAATGAT | TakGATTTGT | CATTTAGCGA | CAATTCTAAG | TCACAAaTAG | AACGTCAACA | 11040 |
| | AACĀĀTTGAG | | | | | | 11050 |
| | • | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 97:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 983 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 97:

CGACATAACG AGGCAAGGGT ACATGATACT TTAGCCTCGT TTTTGATATG TATTTTTCTG 60

AATATAAGGG CAATAGATGG TATTTTATAW TTTTTTTAAG GTAGTGATTA ACATAGATAT 120

55

40

45

| | TCAAGCGGAA | CAGCATTATG | CACCAGTATT | AACGCATTTT | TTAGATCCAA | GAGGGCAATA | 240 |
|----|------------|------------|------------|------------|------------|------------|-----|
| | TATATTGGAA | GTGATTTGTG | GCAGTTATGA | AGATTTAAAC | GTATCTTTTT | ATGGTGGACC | 300 |
| 5 | TAATGCTGAA | AGAAAAAGAG | CAATCATTTC | GCCGAACTAT | TATGAACCTA | AAGAAAGCGA | 360 |
| | CTTTGAATTA | ACTTTAATGG | AAATAGATTA | TCCTGAAAAA | TTCGTCACTT | TAAAACATCA | 420 |
| | ACATATTTTA | GGGACATTAA | TGTCTTTAGG | TATCGAACGC | GAACAAGTTG | GAGATATAAT | 480 |
| 10 | TGTGAATGAA | CGAATTCAAT | TTGTTTTGAC | AAGTAGATTG | GAATCATTTA | TTATGTTAGA | 540 |
| | ATTACAACGT | ATTAAAGGCG | CATCAGTTAA | ACTTTATACT | ATTCCAGTAA | CAGATATGAT | 600 |
| 15 | ACAATCTAAT | GAGAATTGGA | AAAATGAAAG | TGCaCAGTTA | GTTCTTTAAG | GTTAGATGTT | 660 |
| | GTTATTAAAG | AAATGATACG | TAAATCACGT | ACGATTGCGA | AACAACTAAT | ССВАВАВАВА | 720 |
| | CGTGTTAAAG | TGAATCACAC | TATTGTTGAT | TCAGCAGATT | TTCAATTACA | AGCAAATGAT | 780 |
| 20 | TTAATATCCA | TCCAAGGTTT | TGGTAGAGCA | CACATTACTG | ACTTAGGTGG | TAAAACTAAA | 840 |
| | AAAGATAAAA | CGCACATTAC | CTATAGAACA | TTATTCÁAAT | AGTAATGATT | TAAGGAGGAT | 900 |
| | AACAAATGCC | TTTTACACCA | AATGAaATTA | AGAATAAAGA | GTTTTCACGT | GTaAAGAATG | 960 |
| | GTTTTAGAAC | | | | | 2.7 | 983 |
| | | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 98:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10322 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

35

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98:

| | TTTTGCAAAG | CTTATTTTAT | GTCAAACAGA | TAGTCAATGT | GAAACAAAGG | TTAGTACATA | 60 |
|----|------------|------------|------------|------------|------------|------------|-----|
| 40 | TAATCATCCA | GACTTTATGT | ATATATCAAC | AACTGAGAAT | GCAATTAAGA | AAGAACAAGT | 120 |
| | TGAACAACTT | GTGCGTCATA | TGAATCAACT | TCCTATAGAA | AGCACAAATA | AAGTGTACAT | 180 |
| | CATTGAAGAC | TTTGAAAAGT | TAACTGTTCA | AGGGGAAAAC | AGTATCTTGA | AATTTCTTGA | 240 |
| 45 | AGAACCACCG | GACAATACGA | TTGCTATTTT | ATTGTCTACA | AAACCTGAGC | AAATTTTAGA | 300 |
| | CACAATCCAT | TCAAGGTGTC | AGCATGTATA | TTTCAAGCCT | ATTGATAAAG | AAAAGTTTAT | 360 |
| 50 | AAATAGATTA | GTTGAACAAA | ACATGTCTAA | GCCAGTAGCT | GAAATGATTA | GTACTTATAC | 420 |
| | TACGCAAATA | GATAATGCAA | TGGCTTTAAA | TGAAGAATTT | GATTTATTAG | CATTAAGGAA | 480 |
| | ATCAGTTATA | CGTTGGTGTG | AATTGTTGCT | TACTAATAAG | CCAATGGCAC | TTATAGGTAT | 540 |

55

| | GAATGGTTTC | TTCGAAGATA | TCATACATAC | AAAGGTAAAT | GTAGAGGATA | AACAAATATA | 660 |
|------------|------------|------------|------------|------------|---------------------|------------|------|
| | TAGTGATTTA | AAAAATGATA | TTGATCAATA | TGCGCAAAAG | TTGTCGTTTA | ATCAATTAAT | 720 |
| 5 | TITGATGTTT | GATCAACTGA | CGGAAGCACA | TAAGAAATTG | Amtcaaaatg | TAAATCCAAC | 780 |
| | GCTTGTATTT | GAACAAATCG | TAATTAAGGG | TGTGAGTTAG | ATGCCAAATG | TAATAGGTGT | 840 |
| 10 | TCAGTTTCAA | AAAGCGGGAA | AATTAGAATA | TTATACACCT | AATGATATAC | AAGTAGATAT | 900 |
| . • | AGAAGACTGG | GTAGTTGTCG | AATCTAAAAG | AGGCATAGAG | ATAGGTATTG | TTAAAAATCC | 960 |
| | ATTAATGGAT | ATTGCTGAAG | AGGATGTTGT | GTTACCTCTT | AAAAATATTA | TTCGCATTGC | 1020 |
| 15 | TGATGACAAA | GATATTGATA | AATTTAATTG | TAATGAACGA | GATGCTGAAA | ATGCATTAAT | 1080 |
| | ACTATGTAAA | GACATȚGTAA | GAGAACAAGG | TTTGGACATG | CGTTTAGTCA | ATTGCGAATA | 1140 |
| | TACATTAGAT | AAATCGAAAG | TTATTTTTAA | TTTTACGGCG | GATGATCGTA | TTGATTTTAG | 1200 |
| 20 | AAAATTAGTA | AAAATATTAG | CGCAACATTT | AAAAACACGT | ATCGAGTTGA | GACAAATTGG | 1260 |
| | TGTAAGGGAT | GAAGCCAAAT | TGCTTGGCGG | TATCGGACCT | TGTGGTAGGT | CGTTATGTTG | 1320 |
| | TTCTACATTT | TTAGGGGATT | TTGAACCAGT | ATCGATTAAG | ATGGCTAAGG | ATCAAAATTT | 1380 |
| 25 | ATCATTAAAT | CCAACTAAAA | TTTCTGGTGC | ATGTGGTCGT | TTGATGTGTT | GTTTAAAATA | 1440 |
| | TGAAAATGAC | TATTATGAGG | AAGTACGTGC | ACAATTACCT | GATATTGGTG | AAGCAATTGA | 1500 |
| 30 . | AACGCCTGAT | GGTAACGGGA | AAGTAGTTGC | TTTAAATATA | TTAGACATTT | CTATGCAGGT | 1560 |
| | GAAGCTTGAG | GGACATGAAC | AGCCACTTGA | ATATAAATTA | GAAGAAATAG | AAACTATGCA | 1620 |
| | TTAAGGAGGC | ATTATTACAT | TTGGATCGCA | ATGAAATATT | TGAAAAAATA | ATGCGTTTAG | 1680 |
| 35 | AAATGAATGT | CAATCAACTT | TCAAAGGAAA | CTTCAGAATT | AAAGGCACTT | GCAGTTGAAT | 1740 |
| | TAGTAGAAGA | AAATGTAGCG | CTTCAACTTG | AAAATGATAA | TTTGAAAAAG | GTGTTGGGCA | 1800 |
| | ATGATGAACC | AACTACTATT | GATACTGCGA | ATTCAAAACC | AGCAAAAGCT | GTGAAAAAGC | 1860 |
| 10 | CATTACCAAG | TAAAGATAAT | TTGGCTATAT | TGTATGGAGA | AGGATTTCAT | ATTTGTAAAG | 1920 |
| | GCGAATTATT | TGGAAAACAT | CGACATGGTG | AAGATTGTCT | GTTCTGTTTA | GAAGTTTTÄA | 1980 |
| 45 | GTGATTAATC | AAGCACACTC | AAATAGTGTT | ATAATTATAA | atg aatat gg | TTTGGATAAG | 2040 |
| 15 | TCTGAGACAA | TGCATGTTTC | AGGCTTTAAT | TGTGTATAAA | GTTTTGGTGA | TTGCATAAGA | 2100 |
| | GATGGCGGTA | CTAAATGTTA | TTATTAAGTG | TGCACGCAgT | ATCATTAGTT | ATAAAATGTA | 2160 |
| 5 <i>0</i> | GCTGTTAAAA | GTCAAAAATA | CATCGAATGT | AGTTAGGCAT | ATAATATAAA | AAGAGTTTTC | 2220 |
| | AATTACTCAA | TAGAAAAAGG | TTGTCTTCAT | AGGAGTTAAA | aatgttaaaa | GAGAATGAAC | 2280 |
| | GATTTGATCA | ACTAATCAAA | GAAGATTTTA | GTATTATTCA | AAATGATGAT | GTTTTTTCAT | 2340 |

585

| | TGGACTTATG | TTCAGGCAAT | GGGGTGATAC | CCTTGTTATT | GTTTGCGAAA | CATCCACGAC | 246 |
|----------|------------|------------|------------|------------|------------|------------|------|
| | ATATAGAAGG | TGTTGAGATT | CAAAAAACAC | TTGTCGATAT | GGCGCGACGC | ACATTTCAAT | 252 |
| <i>5</i> | TCAATGATGT | TGATGAATAT | TTAACAATGC | ATCACATGGA | TTTGAAAAAC | GTTACTAAAG | 258 |
| | TATTTAAACC | TTCACAATAT | ACTTTAGTAA | CGTGTAATCC | GCCTTATTTT | AAAGAGAATC | 264 |
| 10 | AGCAACACCA | ACATCAAAAA | GAAGCACATA | AGATAGCGAG | ACATGAGATT | ATGTGTACAC | 2700 |
| | TTGAAGATTG | CATGATTGCA | GCCCGTCATT | TATTAAAAGA | AGGTGGCAGG | CTAAACATGG | 276 |
| | TACATCGTGC | AGAGAGACTA | ATGGATGTCT | TGTTTGAAAT | GAGAAAAGTG | AATATTGAAC | 2820 |
| 15 | CTAAGAAAGT | CGTTTTTATA | TATAGTAAAG | TAGGGAAATC | AGCACAAACG | ATAGTAGTAG | 2880 |
| | AAGGTCGAAA | AGGTGGAAAT | CAAGGTTTAG | AAATCATGCC | CCCATTTTAT | ATTTATAATG | 2940 |
| • | AAGATGGTAA | TTATAGCGAA | GAAATGAAGG | AAGTATATTA | TGGATAGTCA | TTTTGTATAT | 3000 |
| 20 | ATTGTAAAAT | GTAGTGATGG | AAGTTTATAT | ACAGGATACG | CTAAAGACGT | TAATGCACGT | 3060 |
| | GTTGAAAAAC | ATAACCGAGG | TCAAGGAGCC | AAATATACGA | AAGTAAGACG | TCCGGTGCAT | 3120 |
| | TTAGTTTATC | AAGAAATGTA | TGAGACAAAG | TCTGAAGCAT | TGAAGCGTGA | ATATGAAATT | 3180 |
| 25 | AAAACTTATA | CCAGACAAAA | GAAATTGCGA | TTAATTAAGG | AGCGATAGTA | TGGCTGTATT | 3240 |
| | ATATTTAGTG | GGCACACCAA | TTGGTAATTT | AGCAGATATT | ACTTATAGAG | CAGTTGATGT | 3300 |
| 30 | ATTGAAACGT | GTTGATATGA | TTGCTTGTGA | AGACACTAGA | GTAACTAGTA | AACTGTGTAA | 3360 |
| | TCATTATGAT | ATTCCAACTC | CATTAAAGTC | ATATCACGAA | CATAACAAGG | ATAAGCAGAC | 3420 |
| | TGCTTTTATC | ATTGAACAGT | TAGAATTAGG | TCTTGACGTT | GCGCTCGTAT | CTGATGCTGG | 3480 |
| 35 | ATTGCCCTTA | ATTAGTGATC | CTGGATACGA | ATTAGTAGTG | GCAGCCaGAG | AAGCTAATAT | 3540 |
| | TAAAGTAGAG | ACTGTGCCTG | GACCTAATGC | TGGGCTGACG | GCTTTGATGG | CTAGTGGATT | 3600 |
| | ACCTTCATAT | GTATATACAT | TTTTAGGATT | TTTGCCACGA | AAAGAGAAAG | AAAAAGTGC | 3660 |
| 10 | TGTATTAGAG | CAACGTATGC | ATGAAAATAG | CACATTAATT | ATATACGAAT | CACCGCATCG | 3720 |
| | TGTGACAGAT | ACATTAAAAA | CAATTGCAAA | GATAGATGCA | ACACGACAAG | TATCACTAGG | 3780 |
| | GCGTGAATTA | ACTAAGAAGT | TCGAACAAAT | TGTAACTGAT | GATGTAACAC | AATTACAAGC | 3840 |
| 15 | ATTGATTCAG | CAAGGCGATG | TACCATTGAA | AGGCGAATTC | GTTATCTTAA | TTGAAGGTGC | 3900 |
| | TAAAGCGAAC | AATGAGATAT | CGTGGTTTGA | TGATTTATCT | ATCAATGAGC | ATGTTGATCA | 3960 |
| 50 | TTATATTCAA | ACTTCACAGA | TGAAACCAAA | ACAAGCTATT | AAAAAAGTTG | CTGAAGAACG | 4020 |
| | ACAACTTAAA | ACGAATGAAG | TATATATAT | TTATCATCAA | ATAAGTTAAT | CACTTTATCG | 4080 |
| | ATTATATGAA | ATTTTAAACG | ATTTTATAAA | CGCAAGCTGT | AATTTTAAAT | GGTAAGTTAT | 4140 |

| | GTITTTTAAT | GTAAAATAAA | TACATTGAAA | GTAATAAATA | CCTTAACATT | GAATAAGATG | 4260 |
|-----|------------|------------|------------|------------|------------|------------|------|
| | AAAATGAGAT | GACGAGATAA | ATGTTCGCGT | CCGTTGAAAT | GCATAGAAAT | CTTAGATATT | 4320 |
| 5 | ATTTGAAGTG | AGACATTACG | AGGAGGAACA | GTTATGGCTA | AAGAAACATT | TTATATAACA | 4380 |
| | ACCCCAATAT | ACTATCCTAG | TGGGAATTTA | CATATAGGAC | ATGCATATTC | TACAGTGGCT | 4440 |
| 0 | GGAGATGTTA | TTGCAAGATA | TAAGAGAATG | CAAGGATATG | ATGTTCGCTA | TTTGACTGGA | 4500 |
| | ACGGATGAAC | ACGGTCAAAA | AATTCAAGAA | AAAGCTCAAA | AAGCTGGTAA | GACAGAAATT | 4560 |
| | GAATATTTGG | ATGAGATGAT | TGCTGGAATT | AAACAATTGT | GGGCTAAGCT | TGAAATTTCA | 4620 |
| 5 · | AATGATGATT | TTATCAGAAC | AACTGAAGAA | CGTCATAAAC | ATGTCGTTGA | GCAAGTGTTT | 4680 |
| | GAACGTTTAT | TAAAGCAAGG | TGATATCTAT | TTAGGTGAAT | ATGAAGGTTG | GTATTCTGTT | 4740 |
| | CCGGATGAAA | CATACTATAC | AGAGTCACAA | TTAGTAGACC | CACAATACGA | AAACGGTAAA | 4800 |
| 0 | ATTATTGGTG | GCAAAAGTCC | AGATTCTGGA | CACGAAGTTG | AACTAGTTAA | AGAAGAAAGT | 4860 |
| | TATTTCTTTA | ATATTAGTAA | ATATACAGAC | CGTTTATTAG | AGTTCTATGA | CCAAAATCCA | 4920 |
| _ | GATTTTATAC | AACCACCATC | AAGAAAAAT | GAAATGATTA | ACAACTTCAT | TAAACCAGGA | 4980 |
| 5 | CTTGCTGATT | TAGCTGTTTC | TCGTACATCA | TTTAACTGGG | GTGTCCATGT | TCCGTCTAAT | 5040 |
| | CCAAAACATG | TTGTTTATGT | TTGGATTGAT | GCGTTAGTTA | ACTATATTTC | AGCATTAGGC | 5100 |
| 0 | TATTTATCAG | ATGATGAGTC | ACTATTTAAC | AAATACTGGC | CAGCAGATAT | TCATTTAATG | 5160 |
| | GCTAAGGAAA | TTGTGCGATT | CCACTCAATT | ATTTGGCCTA | TTTTATTGAT | GGCATTAGAC | 5220 |
| | TTACCGTTAC | CTAAAAAAGT | CTTTGCACAT | GGTTGGATTT | TGATGAAAGA | TGGAAAAATG | 5280 |
| 5 | AGTAAATCTA | AAGGTAATGT | CGTAGACCCT | AATATTTTAA | TTGATCGCTA | TGGTTTAGAT | 5340 |
| | GCTACACGTT | ATTATCTAAT | GCGTGAATTA | CCATTTGGTT | CAGATGGCGT | ATTTACACCT | 5400 |
| | GAAÇCATTTG | TTGAGCGTAC | AAATTTCGAT | CTAGCAAATG | ACTTAGGTAA | CTTAGTAAAC | 5460 |
| 0 | CGTACGATTT | CTATGGTTAA | TAAGTACTTT | GATGGCGAAT | TACCAGCGTA | TCAAGGTCCA | 5520 |
| | CTTCATGAAT | TAGATGAAGA | AATGGAAGCT | ATGGCTTTAG | AAACAGTGAA | AAGCTACACT | 5580 |
| - | GAAAGCATGG | AAAGTTTGCA | ATTTTCTGTG | GCATTATCTA | CGGTATGGAA | GTTTATTAGT | 5640 |
| 5 | AGAACGAATA | agtatattga | CGAAACAACG | CCTTGGGTAT | TAGCTAAGGA | CGATAGCCAA | 5700 |
| | AAAGATATGT | TAGGCAATGT | AATGGCTCAC | TTAGTTGAAA | ATATTCGTTA | TGCAGCTGTA | 5760 |
| :o | TTATTACGTC | CATTCTTAAC | ACATGCGCCG | AAAGAGATTT | TTGAACAATT | GAACATTAAC | 5820 |
| | AATCCTCAAT | TTATGGAATT | TAGTAGTTTA | GAGCAATATG | GTGTGCTTAA | TGAGTCAATT | 5880 |
| | ATGGTTACTG | GGCAACCTAA | ACCTATTTTC | CCAAGATTGG | ATAGCGACGG | Anaattgcat | 5940 |

| | AACCTCAAAT | TGATATTAAA | GACTTTGATA | AAGTTGAAAT | TAAGGCAGCA | ACGATTATTG | 6060 |
|------------|------------|------------|------------|------------|------------|-------------|------|
| | ATGCTGAACA | TGTTAAGAAG | TCAGATAAGC | TTTTAAAAAT | TCAAGTAGAC | TTAGATTCTG | 6120 |
| 5 | AACAAAGACA | AATTGTATCA | GGAATTGCCA | AATTCTATAC | ACCAGATGAT | ATTATTGGTA | 6180 |
| | AAAAAGTAGC | AGTTGTTACT | AACCTGÀAAC | CAGCTAAATT | AATGGGACAA | AAATCTGAAG | 6240 |
| 10 | GTATGATATT | ATCTGCTGAA | AAAGATGGTG | TATTAACCTT | AGTAAGTTTA | CCAAGTGCAA | 6300 |
| | TTCCAAATGG | TGCAGTGATT | AAATAACTGT | ATTTTTAAAA | ATTAGGAGAG | ATAATTATGT | 6360 |
| | TAATCGATAC | ACATGTCCAT | TTAAATGATG | AGCAATACGA | TGATGATTTG | AGTGAAGTGA | 6420 |
| 15 | TTACACGTGc | TAGAGAAGCA | GGTGTTGATC | GTATGTTTGT | AGTTGGTTTT | AACAAATCGA | 6480 |
| | CAATTGAACG | ÇGCGATGAAA | TTAATCGATG | AGTATGATTT | TTTATATGGC | ATTATCGGTT | 6540 |
| | GGCATCCAGT | TGACGCAATT | GATTTTACAG | AAGAACACTT | GGAATGGATT | GAATCTTTAG | 6600 |
| 20 | CTCAGCATCC | AAAAGTGATT | GGTATTGGTG | AAATGGGATT | AGATTATCAC | TGGGATAAAT | 6660 |
| | CTCCTGCAGA | TGTTCAAAAG | GAAGTTTTTA | GAAAGCAAAT | TGCTTTAGCT | AAGCGTTTGA | 6720 |
| | AGTTACCAAT | TATCATTCAT | AACCGTGAAG | CAACTCAAGA | CTGTATCGAT | ATCTTATTGG | 6780 |
| 25 | AGGAGCATGC | TGAAGAGGTA | GGCGGGATTA | TGCATAGCTT | TAGTGGTTCT | CCAGAAATTG | 6840 |
| | CAGATATTGT | AACTAATAAG | CTGAATTTTT | ATATTTCATT | AGGTGGACCT | GTGACATTTA | 6900 |
| 30 | AAAATGCTAA | ACAGCCTAAA | GAAGTTGCTA | AGCATGTGTC | AATGGAGCGT | TTGCTAGTTG | 6960 |
| | AAACCGATGC | ACCGTATCTT | TCGCCACATC | CGTATAGAGG | GAAGCGAAAT | GAACCGGCGA | 7020 |
| | GAGTAACTTT | AGTAGCTGAA | CAAATTGCTG | AATTAAAAGG | CTTATCTTAT | GAAGAAGTGT | 7080 |
| 35 | GCGAACAAAC | AACTAAAAAT | GCAGAGAAAT | TGTTTAATTT | AAATTCATAA | AGTTAAAAGT | 7140 |
| | GAGAAAGATC | ACCGCCATAA | ATGTAAACGA | TGCTATATTC | GTTTAATATG | CTATGGTTCT | 7200 |
| | TTCTCACTTT | AATTAAATTT | AATATCGTGC | ATGTGGAATA | CGTGCGATAG | AGATGGTTAG | 7260 |
| 10 | AGCTTTGAAA | TTAAGAATTG | TAGGAAGGCG | TTTTAAATGA | AAATCAATGA | GTTTATAGTT | 7320 |
| | GTAGAAGGAC | GAGATGATAC | TGAGCGTGTT | AAACGAGCTG | TTGAATGTGA | TACGATTGAA | 7380 |
| 1 5 | ACGAATGGTA | GTGCCATCAA | CGAACAAACT | TTAGAAGTAA | TTAGAAATGC | TCAACAAAGT | 7440 |
| •3 | CGAGGCGTTA | TTGTATTAAC | AGATCCAGAT | TTCCCAGGAG | ataaaattag | AAGTACAATT | 7500 |
| | ACTGAACATG | TCAAAGGTGT | TAAACATGCG | TATATTGATA | GAGAAAAAGC | TAAAAATAA | 7560 |
| 50 | AAAGGGAAAA | TTGGTGTTGA | ACATGCCGAC | TTAATTGATA | TTAAAGAAGC | GTTAATGCAT | 7620 |
| | GTTAGTTCAC | CCTTTGATGA | AGCTTATGAA | TCAATTGATA | AATCTGTGCT | AATAGAGTTG | 7680 |
| | GGGTTAATTG | TTGGGAAAGA | TGCAAGGCGC | CGTAGAGAAA | ТТТТААСТАС | AAAATTYZCZA | 7740 |

| | GCGGATGTAA | GGCAAGCTTT | AGAAGATGAA | TGAGGAAGTG | AAAATGTTGG | ATAATAAAGA | 7860 |
|----|------------|------------|------------|------------|------------|------------|------|
| _ | TATTGCAACA | CCATCAAGAA | CGCGAGCGTT | GTTAGATAAA | TATGGCTTTA | AAAAATTTTA | 7920 |
| 5 | AAGTTTAGGA | CAGAACTTTT | TGATAGATGT | GAATATCATT | AATAATATCA | TTGATGCAAG | 7980 |
| | TGATATTGAT | GCACAAACTG | GGGTGATTGA | AATTGGTCCA | GGCATGGGGT | CATTGACAGA | 8040 |
| 10 | ACAATTGGCC | AGACATGCTA | AAAGAGTATT | GGCATTTGAA | ATTGATCAAC | GTTTAATACC | 8100 |
| | TGTATTAAAT | GATACACTAT | CACCTTATGA | TAATGTGACG | GTGATTAATG | AAGATATTTT | 8160 |
| | AAAAGCGAAT | ATTAAAGAAG | CTGTTGAAAA | TCATTTACAA | GATTGTGAAA | AAATAATGGT | 8220 |
| 5 | TGTTGCAAAC | CTGCCGTACT | ATATTACGAC | GCCAATTTTA | TTAAATTTGA | TGCAACAAGA | 8280 |
| | TATACCAATT | GATGGCTACG | TGGTGATGAT | GCAAAAAGAA | GTGGGCGAAC | GCTTAAATGC | 8340 |
| | TGAAGTAGGT | TCAAAAGCAT | ATGGTTCGTT | ATCAATTGTC | GTACAATACT | ATACAGAGAC | 8400 |
| 20 | TAGTAAAGTA | TTAACGGTAC | CTAAATCTGT | ATTTATGCCA | CCACCTAATG | TTGATTCÄAT | 8460 |
| | AGTTGTAAAA | CTGATGCAGA | GAACTGAACC | GTTAGTAACA | GTAGATAACG | AGGAAGCATT | 8520 |
| | CTTTAAGTTA | GCAAAAGCAG | CATTTGCACA | AAGAAGAAAG | ACAATTAACA | ATAACTATCA | 8580 |
| ?5 | TTTTATTAAA | AAAGATGGTA | AACAACACAA | AGAAGTGATT | TTACAATGGT | TGGAACAAGC | 8640 |
| | AGGTATTGAT | CCAAGACGTC | GCGGTGAAAC | GCTATCTATT | CAAGATTTTG | CTAAATTGTA | 8700 |
| | TGAAGAAAAG | AAAAAATTCC | CTCAATTAGA | AAATTAAATG | ATTGACAAAG | CAAAGCACTA | 8760 |
| | TTGTTAAAAT | TTAAATTTTG | TTTGACGAAA | ACGTTGCAAA | TATGGTATTA | TGTAACTTGT | 8820 |
| | AGCGAGGTGG | AGCAATATGC | CAAAATCAAT | TTTGGACATC | AAAAATTCTA | TTGATTGTCA | 8880 |
| 35 | TGTAGGAAAT | CGTATTGTAC | TGAAaGCCAA | TGGAGGCCGT | AAGA&AACAA | TAAAACGTTC | 8940 |
| | TGGAATTTTA | AAAGAAACAT | ATCCGTCAGT | TTTCATTGTT | GAGTTAGATC | AAGACAAACA | 9000 |
| | CAACTTTGAG | AGAGTATCTT | ATACATACAC | TGATGTGTTA | ACTGAAAATG | TTCAAGTTTC | 9060 |
| 10 | ATTTGAAGAG | GATAATCATC | ACGAATCAAT | TGCACACTAA | ATAAGACATA | TAGAGATGTT | 9120 |
| | AGACGTTTCT | TAGTATAAGA | AGTAAATATT | ATGATAATTA | TTTGAGTGTT | GGGCATTATG | 9180 |
| 15 | TTCAATACTC | TTTTTTTTTA | CAAAATGTTT | AACACTGATG | TTTCGCTTAT | AGATTTTTCA | 9240 |
| ,5 | GTAAATGGAT | AATTGTATTT | ATAAACACAA | ATACAAGTAA | ATACTAAGTA | ATTAGATGGA | 9300 |
| | GAAAATTACT | TTTTTTTAA | AAAAACACTA | AAAAACAAAT | TAAAATGTCA | AATATTAATT | 9360 |
| 50 | CTCTTTATGT | TAAAATCATC | ATATTAAGAT | AACGAAAAGA | GGGCGGAAAA | TGATATATGA | 9420 |
| | AACGGCACCA | GCCAAAATTA | ATTTTACGCT | CGATACACTT | TTTAAAAGAA | ATGATGGCTA | 948 |
| | TCATGAGATT | GAAATGATAA | TGACAACAGT | TGATTTAAAT | GATCGTTTAA | CTTTTCATAA | 9540 |

AAATCTCGCA TATCGTGCAG CGCAACTATT TATTGAGCAA TATCAACTAA AGCAAGGTGT 9660 AACAATTTCT ATCGATAAAG AAATACCTGT TTCTGCTGGC TTAGCTGGAG GTTCGGCTGA 9720 TGCAGCAGCA ACGTTAAGAG GATTGAATCG ACTTTTTGAT ATAGGGGCGA GTTTGGAAGA 9780 ATTGGCTCTA CTAGGCAGTA AAATCGGGAC AGATATTCCG TTTTGTATTT ATAATAAAAC 9840 TGCACTATGT ACTGGAAGAG GAGAGAAAAT CGAGTTTTTA AATAAACCAC CTTCAGCTTG 9900 GGTGATTCTT GCTAAACCAA ACTTAGGCAT ATCATCACCA GATATATTTA AGTTGATTAA 9960 TTTAGATAAG CGTTACGACG TACATACGAA AATGTGTTAT GAGGCCTTAG AAAATCGAGA 10020 TTATCAACAA TTATGTCAAA GTTTGTCTAA TCGATTAGAG CCAATTTCTG TTTCAAAACA 10080 CCCACAAATC GATAAATTAA AAAATAATAT GTTGAAAAGT GGTGCAGATG GTGCGTTAAT 10140 GAGTGGAAGC GGACCTACTG TGTATGGGCT AGCACGAAAA GAAAGCCAAG CAAAAAATAT 10200 TTATAATGCA GTTAACGGTT GTTGTAATGA AGTGTACTTA GTTAGACTAT TAGGATAGAA 10260 GGGTTGAAAA GATGAGATAT AAACGAAGCG AGAGAATTGT TTTTATGACG CAATATTTGA 10320 TG 10322

(2) INFORMATION FOR SEQ ID NO: 99:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5614 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99:

| GATTGATTAA | ATGTTTTAAT | CCACTTCAAT | GCCTTCGATA | AACTCTACAA | TCGCGCTATT | 60 |
|------------|------------|------------|------------|------------|------------|-----|
| CATATAATTA | TTCGATTTCA | TTTGTTCAGC | ATATGTCTCA | TTAAATCCAG | ACATAACTTT | 120 |
| TTTAAAWGCG | AAAATTGAAA | TTGGTATCGT | TACTAATAAG | GCACTAGCCA | TACGCCAATC | 180 |
| AATGAGCATT | ATGTATAAAA | AGATAGCAGC | TGACAAAAGT | AAGTTTCCTA | TAACTTCAGG | 240 |
| AATCATATGT | GCTAAAGGTA | ATTCTATTGT | TTCAACCTTA | TCGACAAATA | TATTTTTTAA | 300 |
| TTCACCTATT | TTCTTAGATT | CCaCTACGCC | TAAÁGGGAGA | CGCATTAATT | TTTGAGCTAA | 360 |
| TTTTTTACGA | ATTTCAGATA | AAATTTCATA | TGCCGTAATA | TGTGATAGCA | TCGTTGACGC | 420 |
| TCCAAAACAA | CACACTTGTG | AAATATAAGC | GATTAAAGCA | ATAAAGATAT | AAACCATAAT | 480 |
| CGAATTAATC | GTATATGTAT | TGTTAATCAT | CATTAAAATA | ATTTTAAATA | CTGCCCAATA | 540 |
| AGGAACTAAT | CCAGAAAAGA | CACTGATGAT | AGACAACAAA | ATTGATAACA | TAATTTTCCA | 600 |

55

5

10

15

20

25

30

40

45

| | ATATGTAACT | CCTKTCAATT | AATAATCTAA | ATTAAGCCGC | TTATATTATT | TATTTCACTG | 720 |
|----|------------|------------|------------|------------|------------|------------|--------------|
| _ | GATGATATAC | ATAATATAAA | TTTGTTATTT | GTTAAAAATT | AATACTTATT | ACAAGTACAT | 780 |
| 5 | CATATATTAG | TTGATAACGA | TTATCAATGT | CGCGTGGATT | TGTGACACAT | TTCTTTTAAA | 840 |
| | AATTCACAAG | GTTATGGGGC | AGAAATGATA | AAGAGCCACT | AATGATTTAT | TATGTAGTGG | 900 |
| 10 | TTCTGGGAGT | GGGACAGAAA | TGATATTTTC | ACAAAATTTA | TTTCGTCGTC | CCACCCCAAC | 960 |
| | TTGCATTGTC | TCTAGAAATT | GGGAATCCAA | TTTCTCTTTG | TTGGGTCCCT | GAATATAGCC | 1020 |
| | TTGTAGAGTC | TAGTACATTG | ATTTGTATCC | CAATGTCCCT | ATAATTGATT | ATTCGCTTTA | 1080 |
| 15 | TCTAATGATC | CTATGACTCA | ACTATTAAAT | CATTTTTCGA | AATACTTAAT | TCTAATATAA | 1140 |
| | TTAAATTCAT | TTATTGTAAT | ATTGCAAAAA | TACATTGCAC | ACCTTGTTCA | TCAATGCTAT | 1200 |
| | AATTAATTAC | ATAATAAATT | GAACATCTAA | ATACACCAAA | TCCCCTCACT | ACTGCCATAG | 1260 |
| 20 | TGAGGGGATT | TATTTAGGTG | TTGGTTATTT | GTCACCTTTT | TTATTGTTGC | GCGTTCGTAA | 1320 |
| | CCAATGTGCA | AAAAACGCAA | CAAGACAGCC | GCTTATAGCT | GAAGTCATGA | TGTTAATTAA | 1380 |
| | TAAATTGAAC | ATCCGTCATA | CACCTCCTCT | CTGCGTTAAA | GTAACGCCCG | AGATGTTAGG | 144 0 |
| ?5 | CGACCATCAT | ATTATATCAT | TTATTTATTA | TATTTCACGC | AATATTAAGG | CTTAAGTAAA | 1500 |
| | GTTTTTTTTA | GTGGTTTACG | CTACTTTAAT | TGCTATCTTT | TAAAATCCAT | TTAGATAATA | 1560 |
| 30 | TAAATGTGAT | GGGTATCGTA | ATAATTAAAC | CAGCAAATGG | TGCAATTTCT | GCTGGCAAAT | 1620 |
| • | TTAGCCAGGA | TACAAATACA | TATAATAAAA | CTGTTTGTAA | GCTTACGTTG | ACAATCTGCG | 1680 |
| | TAATTGGAAA | ACTAATGAAT | TTTCTCCAAG | TAGGTTTTAC | CCTGTAAACA | AAATAACAAT | 1740 |
| 35 | TCAAATAATA | TGAAATCACA | AAAGCGACTA | GAAATCCGGT | AATATGACTA | ATCATATATT | 1800 |
| | CAATGTGTAA | TAATTTTAAC | AGCAATAAAT | AGACAACATA | ATAATTTAAC | GTATTAATGC | 1860 |
| | CGCÇÃACAAT | GATAAATTTT | AAAATTTCAG | CATGCGTTTG | TGTTAGTTTC | ATATGTGTAC | 1920 |
| 10 | TCCTCAACAT | CAAAATATAT | GCATAACTAC | GTTCTCGAAC | ATACTCGAAT | ATGCGAGCCA | 1980 |
| | ATCCGCTTCA | CTTCAAATAT | GCTTATTTCA | ATCTTTATAC | CCTTTCACAG | CAAATTTAGT | 2040 |
| | CTCTTTCCCC | TCATCCTTAT | ACGCCATTAT | AATGTAACTG | ATTTATCGCG | TGACTCATTA | 2100 |
| 15 | GCACTATAGA | GATTACTTTA | GTTCACTAGT | TATATTTTAA | ACAATAAGAG | CGACAACAGT | 2160 |
| | AATGAGAGGA | TGTCTACTAT | GCAATTACAA | AAAATTGTCA | TCGCTCCTGA | CTCATTTAAG | 2220 |
| io | GAAAGTATGA | CCGCACAGCA | AGTTGGCAAT | ATTATAAAAC | AGGCTTTTAC | TAATGTTTAT | 2280 |
| | GGGAATACCC | TTCATTATGA | TATCATTCCG | ATGGCTGATG | GTGGTGAAGG | TACCACAGAT | 2340 |
| | GCTTTAATGC | ATGCAACAGG | TGCCACTAAG | TATACAGTCA | TCGTTAATGA | CCCTTTAATG | 2400 |

| | GCGGCAGCGT | CAGGTTTGGA | TTTATTAGAA | AAAGAGGAAC | GTAATCCTTT | ATACACATCA | 2520 |
|----|-------------|-------------|-------------|-------------|------------|--------------|---------|
| | TCATATGGTA | CCGGTGAACT | AATTAAAGAT | GCATTAAATC | ATGGTGCTAA | GACCATTATT | 2580 |
| 5 | TTAGGGATTG | GTGGCAGTGC | AACAAATGAT | GGTGGTACAG | GTATGCTAAG | TGCACTAGGC | 2640 |
| | GTAAAGTTTA | CTGATGTAAA | CGGGGACTTA | TTACAAATGA | ATGGTGCTAA | TCTTGCTCAC | 2700 |
| | ATTGCACAAA | TCGATATAAC | CAATCTAGAT | TCGCGATTAA | AAGAGGTGAC | CTTTAAAGTG | 2760 |
| 10 | GCCTGTGATG | TTTCAAATCC | TTTATTGGGT | GAAAATGGTG | CTACCTATAT | TTATGGTCCT | 2820 |
| | CAAAAAGGCG | CTGATGCAAA | GATGATACCA | AAGTTGGATT | TCGCAATGTC | GCATTATCAT | 2880 |
| 15 | GATAAGATAA | AAATGTGCAC | AGGAAAGTCC | GTTAATCAAA | TACCAGGTTC | TGGTGCAGCT | 2940 |
| | GGCGGTATGG | GCGCAGCATT | ATTAGCGTTT | TGTGAGACAA | CTTTAACAAA | AGGTATTGAT | 3000 |
| | GTCGTCTTTG | ACATTACAGA | TTTTCATCAA | AGAATTAAAG | ATGCAGACCT | CGTTATTACT | 3060 |
| 20 | GGAGAAGGAC | GCATGGATTA | TCAGACCATC | TTTGGTAAAA | CACCCGTAGG | CGTTGCGTTA | 3120 |
| | GCTGCAAAAC | AATATCATAT | TCCTGTCATC | GCGATTTGTĢ | GCAGTCTAGG | CGAAAATTAT | 3180 |
| | CAACATGTTT | ACGATTTCGG | TATTGATAGT | GCCTATTCTA | TAATCTCTTC | ACCTAGCACT | 3240 |
| 25 | TTAGAAGATG | TCCTACAAAA | TAGCGAACAA | AATTTATTAA | ACACTGCAAC | TGACATTGCT | 3300 |
| | CGTATTCTGA | AATTACAATA | ATGTCAAAGT | AAATCATCAG | CTTTATTATT | TGCAGTTAAA | 3360 |
| | ACTTGAATGA | GGTGAAACCC | ATGAAAAGAA | CTGATAAATA | CCGTGATTCA | TATCAATACG | 3420 |
| 30 | ACAATCAAAA | CCAAAATCAT | CGTCGTCAAT | CTGAAGACGC | ATCGTATAGA | CAACAATATG | 3480 |
| | CTAAAGGCGA | TCCTGAAGAA | CACCCGGAAC | GATACTATAA | TGGTAGAGAT | TATCGAAGAG | 3540 |
| 35 | AACAAATTCT | TGAAGAAGAA | AACGAGAAAT | CCCGCCGTTC | AAAAAAATGG | TTATATATCA | 3600 |
| | TTATTGCCAT | TCTCTTAATT | ATTGTCGCTA | TTTTTGTCAC | ACGCGCCTTA | CTTAACAATG | 3660 |
| | ATAGEGATAA | AGTTAGTAAT | GACCCTAAAG | TCTCTCAAAA | TTATAAAAAA | CAAGTTGAAA , | 3720 |
| 40 | -ATCAAGACGG | -CCAAATTAAC | -CAGCAAGTAG | -ATAATGCTAA | AGAAAATATT | AAAAACAACC | 3.7.8.0 |
| | AAAAAACTGA | TGACATTATT | AAAAATTTAC | AAAATCAAAT | CGACAACTTG | AAGCAGCAAG | 3840 |
| | AACAAAACAA | AGCTGATTCT | AAGCTAACTC | AATTTTATCA | AGACCAAATC | AACAAATTGA | 3900 |
| 45 | CAGAGGCAAA | TAATGCACTT | AAAAACAATG | CAAGCCAAGG | TAAAATTGAA | AGCATGTTAA | 3960 |
| | ATGATATTAA | TACAAAATTC | GACAGTATTA | AATCTAAATT | AGAAAGCTTA | TTTAAAGATG | 4020 |
| | ACAATGGTGG | CGCTAATTAA | TTATTACACC | TGCTTTGATG | ATAAACATTA | ATTCCCTATA | 4080 |
| 50 | CTTTATCTGT | ATCACTACGT | TATTCGTGAT | GATGCATTAA | GAGTATAGGG | ATTTTTTATA | 4140 |
| | TAAACTTGTA | TTCTAACTAC | ATACAAATAC | ACACAAAACG | TATATAATTI | ATATAATTAT | 4200 |

| | IIAIIGCIAA | TTACGTTAGG | CGTCATGACC | GCTTTTGGCC | CACTAACTAT | AGATATGTAC | 4320 |
|-----|------------|------------|------------|------------|------------|------------|------|
| | GTACCATCAT | TACCTAAAGT | GCAAGGTGAT | TTTGGTTCTA | CTACATCAGA | AATTCAATTA | 4380 |
| 5 | ACATTATCAT | TCACAATGAT | TGGTCTTGCA | CTAGGCCAAT | TTATCTTTGG | ACCTTTATCC | 4440 |
| | GATGCTTTTG | GTCGCAAACG | GATTGCTGTA | TCCATTTTGA | TCATTTTCAT | TTTGGTATCA | 4500 |
| . 0 | GGTTTGTCTA | TGTTTGTTGA | TCAATTGCCA | TTATTCTTAA | CTTTACGATT | TATTCAAGGT | 4560 |
| 10 | TTAACTGGTG | GTGGCGTCAT | CGTGATTGCA | AAAGCCTCTG | CTGGTGATAA | ATTTAGTGGC | 4620 |
| | AACGCACTCG | CTAAATTTTT | AGCATCTTTA | ATGGTAGTTA | ATGGCATCAT | CACTATTCTT | 4680 |
| 15 | GCACCATTAG | CCGGTGGATT | AGCTTTATCC | GTAGCAACAT | GGCGTTCTAT | TTTCACAATT | 4740 |
| | TTAACTATTG | TGGCACTCAT | CATTTTAATT | GGCGTCGCTT | CTCAATTACC | TAAAACATCT | 4800 |
| | AAAGATGAAT | TAAAGCAGGT | GAATTTTAGT | AGCGTCATTA | AAGATTTTGG | AAGTCTTTTG | 4860 |
| 20 | AAAAAACCAG | CATTTATTAT | TCCAATGCTA | TTACAAGGWT | TAACTTATGT | AATGCTATTT | 4920 |
| | AGTTATTCAT | CTGCATCGCC | ATTTATTACT | CAAAAATTGT | ATAATATGAC | ACCCCAACAA | 4980 |
| | TTTAGTATCA | TGTTTGCTGT | TAACGGTGTA | GGTTTAATCA | TTGTCAGTCA | AGTCGTTGCT | 5040 |
| 25 | TTATTAGTAG | AAAAATTACA | TCCCCACATA | TTATTAATCA | TTTTAACTAT | TATACAAGTG | 5100 |
| | GTAGGTGTTG | CTTTAATTAT | CCTGACACTT | ACATTCCATT | TACCACTTTG | GGTCTTACTC | 5160 |
| | ATCGCATTCT | TCTTAAATGT | GTGTCCTGTG | ACGTCAATTG | GACCGCTTGG | TTTCACAATG | 5220 |
| 30 | GCTATGGAAG | AACGAACAGG | TGGCAGTGGT | AACGCATCAA | GTTTACTTGG | CTTATTCCAA | 5280 |
| | TTTATCTTAG | GTGGCGCTGT | TGCACCATTA | GTTGGCTTAA | AAGGCGAATT | TAATACATCA | 5340 |
| 35 | CCATATATGA | TTATTATCTT | CATTACAGCC | ATTCTATTAG | TCAGTCTACA | AATCATTTAC | 5400 |
| | TTTAAAATGA | TTAAAAAGCA | ACATGTCGCA | TAACACTTCA | ACATAATTAG | AACCCTAGCA | 5460 |
| | AAGAŦATCTA | TCTTTGTCAG | GGTTCTTCTT | TATGAATTAT | GAGATCGAAT | CTTCAACTAA | 5520 |
| 10 | AATTACGCCT | TCATAGCAAG | GACATTTCTA | TTCAATCACC | CTTTAACAGG | CATCCAAATT | 5580 |
| | TCTGTAATAT | ATTTTTCACT | TGTAGTATCA | CCAT | | | 5614 |

(2) INFORMATION FOR SEQ ID NO: 100:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 9179 base pairs

 - (B) TYPE: nucleic acid (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

50

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 100:

| | AAAGACAATG | ATATGAAGTA | TATGGATATC | ACAGAAAAG | TGCCAATGTC | GGAATCTGAA | 120 |
|-----------|------------|------------|------------|------------|------------|------------|------|
| | GTTAACCAAT | TGCTAAAAGG | TAAGGGGATT | TTAGAAAATC | GAGGGAAAGT | TTTTCTAGAA | 180 |
| 5 | GCTCAAGAAA | AATATGAGGT | TAATGTCATT | TATCTTGTTA | GCCATGCATT | AGTAGAAACA | 240 |
| | GGTAACGGCA | AATCAGAATT | AGCAAAAGGC | ATTAAAGATG | GGAAAAAACG | CTATTACAAC | 300 |
| | TTTTTTGGTA | TAGGAGCATT | CGATAGTAGT | GCTGTTCGTA | GTGGGAAAAG | TTATGCTGAA | 360 |
| 10 | AAGGAACAAT | GGACATCACC | AGATAAGGCG | ATTATTGGTG | GTGCAAAGTT | CATTCGTAAT | 420 |
| | GAATATTTTG | AAAACAATCA | ACTGAATTTA | TATCAAATGC | GATGGAATCC | AGAAAATCCT | 480 |
| 15 | GCGCAACATC | AATATGCGAG | TGACATTCGC | TGGGCAGATA | AAATTGCCAA | ATTAATGGAT | 540 |
| | AAATCCTATA | AGCAGTTTGG | TATAAAGAAA | GATGATATTA | GACAAACATA | TTATAAATAA | 600 |
| | GACATCGGTG | CTTAAAGGAG | CTGGAACAAT | TTATTGTTTC | GAGCTCCTTT | AGCGCATTCT | 660 |
| 20 | GAGTGTGTTA | GTTAAATGGA | TTTTAACCTA | ACAAAAAACG | CTATATAGCA | TCAAATATGC | 720 |
| | TATATCCCAC | ATCATTGTTA | CAAATGTACA | TGATGTAAAT | GAATATTGCT | GTCTAAATGT | 780 |
| | GCATGTAATA | TACAATGGTG | CAGATAATAC | ACTTAAGTCC | TTAAAAATGA | AACGTTAgTT | 840 |
| 25 | CCAAGAGTCA | TTTTTAAACA | ATAGTGCATG | TGATAAAATA | GAAAAGAATG | AAAAATATAG | 900 |
| | AGGTGACAAT | ATGAAGATAG | CAATTATAGG | TGCAGGCATC | GGTGGATTAA | CAGCTGCTGC | 960 |
| | ATTATTACAA | GAACAAGGTC | ATACTATTAA | AGTCTTTGAA | AAAAATGAGT | CAGTTAAAGA | 1020 |
| <i>30</i> | AATTGGCGCT | GGGATTGGTA | TCGGAGATAA | TGTGCTTAAA | AAACTAGGTA | ATCATGACTT | 1080 |
| | AGCTAAAGGT | ATTAAAAATG | CTGGGCAAAT | CTTATCTACA | ATGACAGTGT | TAGATGACAA | 1140 |
| 35 | AGATCGCCTG | TTAACTACTG | TTAAATTAAA | AAGTAATACA | TTGAATGTGA | CGTTACCACG | 1200 |
| 33 | CCAAACATTA | ATTGACATTA | TTAAATCTTA | TGTAAAAGAT | GACGCAATAT | TTACAAATCA | 1260 |
| | TGAAĢŦCACG | CATATAGATA | ATGAGACAGA | TAAAGTTACC | ATACATTTCG | CGGAACAAGA | 1320 |
| 40 | AAGTGAAGCA | TTTGATTTAT | GTATTGGTGC | TGATGGAATT | CATTCTAAAG | TGAGACAATC | 1380 |
| | TGTAAATGCT | GACAGTAAAG | TATTATATCA | AGGGTATACA | TGCTTTAGAG | GTTTAATTGA | 1440 |
| | TGATATTGAT | TTAAAGCATC | CGGaTTGTGC | AAAAGAATAC | TGGGGaAGAA | AAGGAAGAGT | 1500 |
| 45 | AGGTATTGTT | CCGTTATTAA | ATAATCAAGC | ATATTGGTTC | ATTACAATTA | ACTCGAAGGA | 1560 |
| | AAACAATCAT | AAATATAGTT | CGTTTGGTAA | ACCTCATTTG | CAAGCATACT | TTAATCACTA | 1620 |
| | TCCAAATGAA | GTTAGAGAGA | TCTTAGACAA | ACAAAGTGAA | ACAGGTATCT | TATTGCATAA | 1680 |
| 50 | TATTTATGAT | TTGAAACCAC | TCAAATCTTT | TGTTTATGGT | CGTACTATTT | TACTAGGAGA | 1740 |
| | TGCAGCACAT | GCGACAACGC | CTAATATGGG | GCAAGGTGCT | GGACAAGCAA | TGGAAGATGC | 1800 |

| | TAAAATACGT | GTCAAACATA | CTGCAAAAGT | AATTAAGCGT | TCTAGAAAAA | TCGGTAAAAT | 1920 |
|----|------------|------------|------------|------------|------------|-------------|------|
| | TGCCCAATAT | CGTAGTCGTT | TATTTGTTGC | AGTTAGAAAT | CGTATTATGA | AAATGATGCC | 1980 |
| 5 | AAATGCATTA | GCAGCTGGAC | AAACTAAATT | СТТАТАТААА | TCGAAAGAAA | AATAATACAA | 2040 |
| | CAATATGAAA | ACCCCCGTAT | GTTGAAACGA | GAGCTCAACA | TATGGGGGTT | CTTGTTTTTA | 2100 |
| 10 | TAATGTTATT | ATAATAAATT | CAATTATTAG | TTAACGACAA | ATTGTGGTTT | CTCACCTTGA | 2160 |
| | ACGGCACTAA | TTGCAGCATT | AGCAACAATT | TTAGACATCA | TGTCACGTGC | TTCAAATGTA | 2220 |
| | GCATTACCAA | TATGCGGTGT | TAATACTACA | TTATTAAGTG | ATTTTAAGTC | ATCGGTAATA | 2280 |
| 15 | TCTGGTTCAA | ATTCATATAC | ATCAAGTGCA | GCACCTTCAA | TTTCATTATC | TTTCAATGCT | 2340 |
| | TGCACTAGTG | CTTGTTCGTG | CACGATTGGA | CCACGAGAGG | CATTGATTAA | ATACGCCGTA | 2400 |
| | GATTTCATCA | TTTTAAATTG | TTCTGTATCA | ATTAAATGAT | GCATTTTAGG | ATTATAAGCA | 2460 |
| 20 | GCGTTGATAG | TGATAAAATC | TGCATTCTTT | AATAGTGTAT | CTAAATCTAC | ATATTTTGCA | 2520 |
| | CCGATTTCTC | GTTCTTTTTC | TTCTTTGCGA | TTAGGTCEAG | TGTATAGCAC | ATCCATGTCA | 2580 |
| | AATGCTCTTG | CACGACGAGC | TACTGCACTA | CCAATTTCAC | CTAAACCGAT | AATGCCGATT | 2640 |
| 25 | GTTTTCCCAG | ATACTTCTCT | ACCTCTGAAA | AATAAAGGTG | CCCATCCATC | AAATCCAGTT | 2700 |
| | GTACGTGATA | ATTGGTCCCC | TTCAACAATA | CGACGCGCTA | CTGCAAGTAC | TAATCCAATT | 2760 |
| 30 | GTTAAATCAG | CAGTCGCGTT | TGTTGATGCT | TTAGGTGTGT | TTGTAACATC | TATACTTTTT | 2820 |
| | TCTCGGGCAT | ACTCGATATC | AATATTATTA | AAACCAGCGC | CATAGTTGGC | AATGATTTTT | 2880 |
| | AAGTCTTTAC | CAGCATCGAT | AACATCTTTA | TCAACGTTTG | TAGATAATAA | ACTAATTAAG | 2940 |
| 35 | GCAGTCGCGT | TTTTAACACC | TTTAATTAAA | GTGTCTTTAT | CGACTAATCC | TTTACCTTCA | 3000 |
| | TACATTTCAA | CTTCAAAATG | TTCTTGTAAA | AGTTTTAAAC | CTACTTCTGG | TATEGCACCA | 3060 |
| | gCAACATAAm | CTTTTtCCAT | AAAAGAtCAC | TCCTTTTATC | TTAGTATAGT | AGAAGATTAG: | 3120 |
| 40 | ACAGTATACA | ACTATGTCAT | GATGTCTTGT | GTATCAATGA | TGTAAGCGCG | TACTITIGAT | 3180 |
| | GGAGGCGATA | TAACTTAGGC | ACTGTAGAAC | TATGAATATT | GTAATGTGGA | AAAACTGGAT | 3240 |
| | CAATTAAATT | AGATAACGTA | GTTTTAAAGT | TAATAGTATT | AGAAAAATT | AATATTTTGA | 3300 |
| 45 | ATATGGGAGG | TAAATATAAA | AAGTAGGTGG | CAACGAAAAA | TAGCAAAAA | AGAGCTTCTC | 3360 |
| | CTATAAAGGA | AAGCTCAAAG | TTTTTTGATG | ACATATGTAC | TAGAATTAAG | TTTCAAGACA | 3420 |
| 50 | ATATGTATCA | TCGTGTTTAT | ATTAAATATG | GATGTAGTTG | TAGTTACCTG | CTTCACTTGC | 3480 |
| | AGAAATAGTT | CTAGAACTTA | CTGAGAAAGG | TCCGCCACTA | TAATTCATTT | CTGAAATTGT | 3540 |
| | AACTGAACCA | TCACTGTTTA | CACTITCTAC | ATATGCAACG | TGACCAAATG | GTCCTTCAGA | 3600 |

| | AGCAGCAGCC | CAATTATTAG | CATTTCCCCA | AGTAGAACCG | ATTTCTCCGC | CAACTTTATC | 3720 |
|----|------------|------------|--|-----------------|--------------------|--|------|
| | ATATACATAC | CAAGTACATT | GTCCTGCAGT | GTATAAGTTA | CCAGAATGTG | AAATTGATGA | 3780 |
| 5 | TGTAGTTGTC | GTAGTTGTCG | TAGTCGTTGT | AGTTTGAGTC | GTGTTGTAGT | TATAGTTGTT | 3840 |
| | GTAATTTGTA | TAATTTTCAG | CAGCATCTGC | ATGATGTGCT | TGACCTACTA | ATGCTGTGCC | 3900 |
| | GATTCCTGCT | GTTAACGTAG | TTGCTGTTAC | TAATTTTTTC | ATGAATAAAG | TCCTCCAAAG | 3960 |
| 10 | TTCTATATCT | TTTTTTATAA | ATAAAACGTA | GCGACTGTTT | TATTCTCACA | TCTCGAATTG | 4020 |
| | ATGACAATAG | TTACTTTAAC | AAAATtAATG | cttcttgtgg | GGAATGTTAT | TGATTTGTAA | 4080 |
| 15 | aagaataaaa | AAACTTTGAC | TAATTTTGTA | ATAAAAATTA | GTCAAAGTTA | CAATGAGATT | 4140 |
| • | AACAGATAAT | TAATAGGAAA | TATTTATTTG | TAATATGTTT | AAATAAATCG | AATTGTTAAA | 4200 |
| | GGTATTATAT | ATTCTTGGCC | ATTATAATAT | TTGACACACG | CAATAATTGT | GAATACAAAA | 4260 |
| 20 | GATAATATTG | AGAAAGCGAA | TATGGATAAA | ATACCGATAA | ACGTAATGAT | GAAACCTATA | 4320 |
| | ATAATAATGA | AATCAATATC | TGTAGCAATT | AGGAAAACGC | CTATTAAAGT | GATAACGACT | 4380 |
| | AAAACGATAG | ACCAAATAAT | ATAAGAAATC | GTATAGTTAA | GATAATTITT | TCCAGCACGA | 4440 |
| 25 | TCAACTAGTT | TCGATTCATC | TTTTTTCAAT | AACCATATTA | TCAGTGGACC | AATAATAGAT | 4500 |
| | GTGAATAAAC | TTAATAAATA | GATAAGCATC | GCCATAATGT | TCTCATCATT | GGATTTGCGA | 4560 |
| | TTCGGTTGAT | GATTTGTTAC | GTCGTTCATT | TCAGTTGTCA | TATTAGACAC | TCCTTTGAAA | 4620 |
| 30 | ATTGTAATAT | TATCTTTAAC | TATAACAAAA | TATAATCAAA | AATAAACATG | TTTATTAAAC | 4680 |
| | AATTATTAAA | ATAAAAATA | ATTGGTGGAC | GTCGGCGTTT | AAATAGGTTA | ATTTAAGGTT | 4740 |
| 35 | ATATATACTT | AACATTTATA | ATGATGCGTA | ATGAATTCGC | ATCATTTTTA | TATTGTCTTA | 4800 |
| | CGTATAATTT | GTTTTTAATT | TTAACCAAAG | ATAGAAAGAG | GGTTGTTTAT | GAAAATAGCA | 4860 |
| | ATTGTAGGAT | CAGGAAATGG | CGCAGTTACG | GCAGCAGTAG | ATATGGTGAG | CAAAGGCCAC | 4920 |
| 40 | GATGTTAAAT | TATATTGTCG | TAATCAATCT | ATAAGTAAGT | TTCAAAACGC | AATCGAAAAG | 4980 |
| | GGCGGATTTG | ATTTTAATAA | TGAAGGTGAT | GAACGTTTCG | TAAAATTCAC | TGATATTAGT | 5040 |
| | GATGATATGG | AATATGTTTT | AAAAGATGCT | GAAATTGTTC | AAGTGATTAT | TCCATCTTCA | 5100 |
| 45 | TACATAGAGT | ATTATGCTGA | TGTAATGGCA | GAGCATGTAA | CTGATAATCA | GTTGATATTC | 5160 |
| | TTCAACATGG | CTGCAGCAAT | GGGGTCAATT | CGTTTTATGA | ATGTTTTAGA | AGATAGACAT | 5220 |
| | ATTGAAACAA | AACCACAACT | AGCGGAAgcT | AATACGTTGA | CGTATGGTAC | GCGTGTCGAT | 5280 |
| 50 | TTTGAAAATG | CAGCAGTTGA | TTTATCTCTA | AATGTACGTC | GTATCTTCTT | TTCAACATAT | 5340 |
| | CATAGRAGOT | CTCTAAATCA | ייים אייים איים אריים איים איים איים איי | A A A COTTO A A | מביצה מידיים הידים | でいる イブランス イブ・イブ・イブ・イブ・イブ・イブ・イブ・イブ・イブ・イブ・イブ・イブ・イブ・イ | 5400 |

| | CCAACATTAT | TGAATGTCGG | TCGCATTGAT | TATGCTGGCG | AGTTCGCTTT | ATATAAAGAA | 5520 |
|----|--------------|--------------|------------|------------|------------|------------|------|
| | GGAATTACTA | AACATACAGT | TAGATTACTI | CATGCAATCG | AATTAGAACG | TTTGAATTTA | 5580 |
| 5 | GGCCGTAGAT | TAGGTTTTGA | ATTATCAACA | GCTAAAGAAT | CACGTATTGA | ACGTGGTTAT | 5640 |
| | TTAGAACGTG | ATAAAGAAGA | TGAACCATTA | AATCGTTTGT | TTAATACAAG | CCCAGTATTT | 5700 |
| 10 | TCACAAATTC | CAGGACCAAA | TCATGTAGAA | AGCAGATATT | TAACTGAAGA | TATTGCATAT | 5760 |
| 70 | GGTTTAGTAC | TATGGTCAAG | CTTAGGTCGT | GTTATTGATG | TACCGACACC | AAATATAGAT | 5820 |
| | GCAGTAATTG | TAATTGCATC | AACCATTTTA | GAGAGAGACT | TCTTTGAGGA | AGGCTTAACA | 5880 |
| 15 | GTTGAAGAAA | TTGGTTTAGA | TAAGCTTGAT | TTAGAAAAAT | ATTTAAAATA | AATGATGGCT | 5940 |
| | TGAAGATAGA | AAAGGATATA | GCATTATGCA | AAAGCAATAA | ATTGAAGAAA | AGAGGTTTCT | 6000 |
| | CATCAATAAG | CGnAGGGGAC | GATAGATGAT | GAAAAGAAAA | CCCACCTTTT | TAGAATCAAT | 6060 |
| 20 | TTCGACAATG | ATTGTAATGG | TTATTGTTGT | TGTAACAGGC | TTTGTGTTTT | TTGATATTCC | 6120 |
| | AATTCAAGTA | TTATTAATTA | TIGCCTCAGC | ATATGCCACA | TGGATTGCAA | AACGTGTAGG | 6180 |
| | CTTAACATGG | CAAGATTTAG | AAAAAGGCAT | TGCAGAACGT | TTAAATACTG | CAATGCCTGC | 6240 |
| 25 | AATTTTAATT | ATACTAGCGG | TAGGAATTAT | AGTAGGCAGT | TGGATGTTTT | CTGGCACAGT | 6300 |
| | GCCAGCCTTG | ATTTATTATG | GCTTAGATTT | ATTGAATCCA | AGCTATTTT | TAATATCAGC | 6360 |
| | CTTTTTTATA | AGTGCTGTTA | CATCTGTAGC | AACTGGTACA | GCATGGGGCT | CTGCATCAAC | 6420 |
| 30 | TGCAGGGATT | GCACTTATTT | CTATTGGTAA | TCAATTGGGG | ATTCCTCCAG | GGATGGCAGC | 6480 |
| | GGGTGCTATT | ATAGCAGGGG | CTGTGTTTGG | CGATAAAATG | TCACCATTAT | CAGATACAAC | 6540 |
| 35 | TAATTTAGCG | GCGCTTGTTA | CTAAAGTTAA | AAATTTATAT | CATATACATT | CGATGATGTG | 660Ö |
| | GACGACGATA | CCTGCATCAA | TCATAGGTTT | ATTAGTATGG | TTTATTGCTG | GATTTCAATT | 6660 |
| | TAAAGGGCAT | TCAAATGATA | AACAGATTCA | AACTTTGTTA | TCAGAGCTTG | CACAGATTTA | 6720 |
| 40 | TCAAATTAAC | ATATGGGTCT | GGGTTCCCTT | AATTGTGATC | ATTGTTTGTT | TGCTATTTAA | 6780 |
| | AATGGCTACA | GTGCCAGCTA | TGCTAATATC | AAGCTTTTCT | GCCATTATAG | TGGGGACTTT | 6840 |
| | TAATCATCAT | TTCAAAATGA | CAGATGGTTT | CAAAGCAACA | TTTAGTGGTT | TTAACGAATC | 6900 |
| 45 | AATGATACAT | CAGTCTCATA | TTTCATCCAG | TGTGAAAAGC | TTGTTAGAAC | AGGGTGGTAT | 6960 |
| | GATGAGTATG | ACCCAAATAT | TAGTAACGAT | ATTTTGCGGA | TATGCATTTG | CAGGTATTGT | 7020 |
| | AGAAAAAGCA | GGATGTTTAG . | AAGTCTTATT | AACTACTATT | TCTAAAGGCA | TCCATTCTGT | 7080 |
| 50 | AGGAAGTTTA A | ATATGTATTA | CTGTTATTTG | TTGTATTGCG | CTTGTATTCG | CTGCAGGTGT | 7140 |
| | TGCTTCGATT | GTAATTATTA ' | TGGTCGGTGT | GTTAATGAAA | GATTTGTTCG | AAAAATACCA | 7200 |

| | AATACCATGG | GGAACATCAG | GTATTTACTA | TACGAATCAA | CTTCATGTCT | CTGTTGAAGA | 7320 |
|----|------------|------------|------------|------------|------------|------------|------|
| | ATTTTTCATA | TGGACAGTAC | CATGTTATTT | ATGCGCAATT | ATAGCAATTA | TCTATGGTTT | 7380 |
| 5 | TACAGGGATA | GGTATTAAAA | AGTCATCGAA | TTCACGTTTA | ACTTAATGTG | AGCGTGGAAT | 7440 |
| | ATATATAATA | TGTTGAAACA | CTTTAATCAT | TTATAATTGT | AGCGGTTATA | ATTTGAAAAG | 7500 |
| | GTTTTAACTT | AGAATAAATA | TCCTCTATGC | ATATACTGAA | TATGTTTTGT | AGCGGAACAT | 7560 |
| 10 | GTTGATATAT | GTAATGTAAG | TTTTATGTCA | TGATTTGTAA | TGACTAAATT | AATTGAGAAT | 7620 |
| | TTGAAGGCAA | GTATATTTGT | AAGTACTTTA | ACTAAAAATT | TATCAATGTA | TAGCCGATTT | 7680 |
| 15 | GACATGCCTA | AATTTGGGTG | TGTCAATGGC | TGTATGTTGT | TTATTCTTTA | TTACAGAGTG | 7740 |
| | AATCGGATTG | GTGAAAATCG | AAATTTTGAG | ATTTTTACCA | ATTCGATTTT | TTTCATAGAA | 7800 |
| | ATTAAAAAAG | CCAACAAGGC | TCTTGAAACC | TTGTTGGCGT | AAACATAGCC | ATCACTAATT | 7860 |
| 20 | AGTGAATGAA | GTTATAACCA | GCAGCTTGGC | TAGCTGAGAT | TGTACGTGAA | GTTACAACAC | 7920 |
| • | CTGGGCCATA | ACCATAGTTC | ATTTCTGAAA | CTCTTACTGA | ACCATTGCTG | TTAACACTTT | 7980 |
| | CAACGTATGC | AACGTGACCG | TATGCACCTT | GAGTTGTTTG | CATAATTGCA | CCAGCTTTTG | 8040 |
| 25 | GTGTATTGTT | CACTGTGTAA | CCAGCTCTTG | CAGCTGCGTT | AGCCCAGTTA | CTTGCATTGC | 8100 |
| | CCCAAGTTGA | ACCGATTTTA | CCACCTACAC | GATCAAATAC | GTAGTATGTA | CATTGACCAG | 8160 |
| 20 | AAGTGTATAA | GTTACGTCCT | GAAGTATAAC | CACTTGAGAT | TGAACGGCCA | TTTGATGATG | 8220 |
| 30 | GAGCCATAGT | TGTAGTTACT | TGAACATTGT | TGCTTGAAGT | GCTGTAGCTT | GCACCTAAAC | 8280 |
| | CACCAGTACG | GTAGCTGTTT | GTGTTGTAAC | TATTATAGTT | ATTGTAGTTA | TATGATTGAT | 8340 |
| 35 | TATTATTTGA | GTAGTTGTTG | TAACGGCTGT | AGTTATTGTA | GCTATAACCG | TTGTTGTAAT | 8400 |
| | TGTTATAGTT | ATTGTAACCA | TTGTAGTAGT | AATAGCTGTA | GTAGCCATTA | TCTTGGTTTA | 8460 |
| | ATTGACTTGG | ATGCCAGTTA | CCTTTCCATG | TGTAATGGTA | GTTACCTTGT | GCATCAATAG | 8520 |
| 40 | TGTAAGTATA | GCTATATGAT | GTTGGGTCGT | TTGGATTATA | ACCGTAGTTA | TCTTGCTCAG | 8580 |
| | AAGCATGAGC | TTGATTTCCT | GATGCAATTG | CGATTGTAGC | GAATCCTGCA | GTTGCGATAG | 8640 |
| | TAGCTGTAGC | GATTTTCTTC | ATTTTAAAAA | TATCCTCCTA | AATTTTAA | ATCTAAAATA | 8700 |
| 45 | TTTTCGTAAT | GTCCGTGTGA | CAAAATTAAT | GTTATAAGTT | ATCTCTCGTA | ATTAAACGAC | 8760 |
| | AAGAAAGACT | ATAACAGAAA | TTAGCGTCCT | TGTGTGCTTT | GTTAACGTTT | TGTAATTTTT | 8820 |
| 50 | TGCTAATATC | TTGACACAAT | AGAATTTTAA | AAGTATAGAA | ATTTGCATTT | TGCAAAACTT | 8880 |
| 50 | ATAACTACGG | CATTCTTTGT | GAAAACTGAA | TGTTTCGAAA | ATAAGTCTGT | TACAAATTTG | 8940 |
| | TAATATTACT | GAAAATTCTA | AATGTATATT | TTGTGCATAA | TATAGGACTT | ТТААТСАСАА | 9000 |

| | GGATGAAAAT GTATATTTAA TGGATAAAAT ATCCTAATTT AGCATAAAAA AATGTTTTAA | 9120 |
|-----|--|------|
| | TAAAAGTATT ATTTGATATA ATCGATTTAT GTTTTGTTAC TGCTAAAAAA CATGTGGCG | 9179 |
| 5 | (2) INFORMATION FOR SEQ ID NO: 101: | |
| | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1868 base pairs (B) TYPE: nucleic acid | |
| 10 | (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 15 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 101: | |
| ,,, | CCTTCAGCCA TTTGACTTCG ACATGAGTTG CCTGTACATA TAAAATAAAT TGTTTTTTTTA | 60 |
| | GTCATAACAA TCTCCTAATT AATTAAAATA TGATAAGTGT TAGATACAAC CCTATGAGGG | 120 |
| 20 | TTATAAATAG TACTGGAATT GTAATGATGA TACCAGTTTT AAAGTATGTG CCCCAAGAAA | 180 |
| | TCTTAACATC TTTTTGtGTT AAGACGTGTA ACCACAGTAA TGTAGCTAAA GAGCCTATCG | 240 |
| | GTGTAATTTT TGGACCTAAA TCAGAACCGA TAACATTCGC ATAAATTAGG CCTTCTTTTA | 300 |
| 25 | ACATGCCATG GACATTTGAT TGACCAATAG CAATCGCATC TATTAAAACT GTAGGCATAT | 360 |
| | TATTCATTAT TGATGATAAA AACGCTGAAA TGAAGCCCAT TCCCAAAATA GTGCTAAATA | 420 |
| | GACCGTAATT GGAAATATAT TCTAATATTT TAGCCAATAT TAAAGTAATG CCAGCATTTC | 480 |
| 30 | TTAAGCCGAA TACGACGATA TACATACCAA TTGAAAATAA TACTATATTC CAAGGTGCGC | 540 |
| | CCTTAATGAC TTGCTTAATA TTTACAGCAT TTGATTTACG AGCCAACATT AGAAAAATAA | 600 |
| 35 | AAGCAATGAT TCCAGTGAAA ATTGATACCG GAATTTTAGT AAATTTACTG ATTAGATAGC | 660 |
| | CGAAAAGTAA TATAACTAGA ACAATCCATG AAATTTTAAA TAGCTTTAAA TCATTAATGG | 720 |
| | CATCETTAGG ATGCTTTATA TTATTATCAT CAAACGTTTT AGGTATCGCT TTTCTAAAAT | 780 |
| 40 | ATAACCACAA TACTATAATA CTTGCTAAAA GCGAGAATAA ATTAGGTATA ATCATTCTAC | 840 |
| | TAAAATATCG AACGAATCCT ACATGAAAAT AATCAGCAGA TATAATATTC ACTAGATTGC | 900 |
| | TCACGATTAA AGGTAAAGAA GTTGTGTCAG CTATAAAACC ACTCGCAATA ATNAAAGGGA | 960 |
| 45 | ATATGGCCCG CTTACTAAAA CCTATATTTT TAACCATCGC TAATACAATA GGCGTTAAGA | 1020 |
| | TTAACGTGCG CCATCATTTG CGAAAAATGC AGCAACAATG GCACCCAATA ATATGATATA | 1080 |
| | AACGAACATT TTTAAACCAT TGCCTTTTGA AGCATGAAGC ATGTGAATAG CTGACCATTC | 1140 |
| 50 | GAATAATCCA ACTTTATCTA ATATTAATGA AATAAGAATG ACTGAGACAA AAGTCAAAGT | 1200 |
| | AGCATTCCAA ACAATACCTG TTACTTCGAA AACATCGGAA AAACTTACAA CACCAGTAAT | 1260 |

| T | AATACAAAT | AATAAAGTTA | CTAGAAAAAT | GAGTGTCGCT | AAAGTTGTCA | TCATTAGCAT | 1380 |
|----|-----------|------------|------------|------------|------------|------------|------|
| T | CACCAGTCT | TAAGGTTATG | ACAAATACAT | CGTTGGTTAG | AGGTATGAAC | CTTAGACAAG | 1440 |
| T | ATTAATTA | CGGACTCAAA | AATATTATGA | TTgAGCTGGT | ATAAATGTTT | ATTTCCGATT | 1500 |
| T | TCGTGTCG | TAACTAAGTT | GGTTTTTACT | AATGCTTTCA | TATGTTAGCT | AAGTGTAGGT | 1560 |
| TO | SAGAGAATT | GAAAATGTGC | TAACAAATCA | CAAGCGCATA | ACTCTCCACA | AGAAAGTAAA | 1620 |
| TC | TAGTATTT | CTAATCTGCT | TGAATCTGAT | AAAACTTTTA | AAAATGTTGĊ | TAGTTCTTTA | 1680 |
| TF | CGTCATAA | CATACCTCCT | AGACGTTAAA | TAGATTATCA | TCTATATAGA | TGAATGTCTA | 1740 |
| TC | TTCCTTTG | GTATATTACA | CGATATGACT | ATGTAATTTA | AATTTGGTTT | TAGTATTAAA | 1800 |
| AG | GGTATTAA | AGATAAATTA | TAGATATTGA | TTTTGCAAAA | TATACTCTTT | GTTCTGCATT | 1860 |
| GA | AAAAGG | | | | | | 1868 |

(2) INFORMATION FOR SEQ ID NO: 102:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15249 base pairs
- (B) TYPE: nucleic acid (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 102:

| ATTTATGAAA TCCATA | AGCNA TAAACATTAT | TCTTGCATCG | GCTATACAAA | CAGTTACCGC | 60 |
|-------------------|------------------|-------------|------------|------------|-----|
| AAGCAAATTT GTATAT | CAAC CTGGAATTGT | GTTCACGTCA | ATGGCaAATG | CCGATGATGT | 120 |
| GTTATCAGGC GATAGT | TTATT TTATGGCTGA | ATTAAAATCT | ATTAAGCGTA | TTGTTGAAAT | 180 |
| TCCAGATAAT CAAAAA | AATAT ACTGCTTTAT | AGATGAAATT | TTTAAAGGTA | CCAACACAAC | 240 |
| TGAACGAATT GCCGCT | TCAG AATCAGTACT | ATCATTTTTA | CATGAAAAAT | CTAACTTTAG | 300 |
| AGTTATTGCA GCAACA | CATG ATATTGAGTT | AGCTGAACTC | TTAAAACAAC | GTTATGAAAA | 360 |
| TTACCATTTC AATGAG | GTAA TAGAAAATAA | TAACATACAT | TTTGATTACA | AAATTAAGCC | 420 |
| TGGCAAAGCA AATACA | CGTA ATGCCATCGA | ATTATTAAAA | ATCACTTCAT | TTCCAGCAAA | 480 |
| AATATATGAA CGAGCA | AAAG ATAATGTCCC | GAAAATTTAG | CATTTAACTT | TAAACATAAA | 540 |
| AACGTCAGCT ATCACA | TGAC AGAAGACTAT | GAACAGTTTC | AATAATGTTC | ATAGTAATCA | 600 |
| TGTTAATAAC TGACGT | TTAT TTTATTCTGC | AGAATACTCT. | TCTAAATCTA | TATTGCTGTG | 660 |
| CCCATTTAAT GCTAAA | TCAG CAAATCGACC | TTGCTGATAC | AAATAGTGGC | CGGCAACGCC | 720 |
| TATCATTGCA GCATTA | TCTG TGCATAATTT | AGGACTTGGG | ATAGTTAATT | GAATGTCATT | 780 |

55

5

10

15

20

25

35

40

45

| | AACAATTAAT | CGCTGAACAC | CATATTCTTT | ACAAGCTTGA | ATAGCTTTAA | ACGTGAGCAC | 900 |
|----|------------|------------|------------|------------|------------|-----------------|------|
| _ | CTCTACAACA | CTGTTTTGAA | AGCTCGTTGC | TACGTTAGCT | TCAATGATTG | Gaatatttt | 960 |
| 5 | TTGTCGTTGA | TTGTGAAGTT | GATTGATTAC | GGCACTTTTC | AACCCACTAA | AACTAAAATC | 1020 |
| | ATAACTATCT | TTATCCAACC | AAACACGAGG | GAATGAATAA | GTATCTTCAC | CTTCAGCAGC | 1080 |
| 10 | CAACCGATCA | ACTTGTGGAC | CACCTGGATA | ATTTAAACCA | ATTGTTCGTG | CCACTTTATC | 1140 |
| | ATAAGCCTCA | CCTACTGCGT | CATCTCGTGT | TTCACCAATG | ACTTCAAATG | ATAAATGATC | 1200 |
| | CTTCATATAA | ACTAATTCAG | TATGTCCACC | TGAAACAATA | AGTGCAATTA | GCGGGAATGT | 1260 |
| 15 | TAATGGCTCT | TCTATGTGAT | TAGCATATAT | ATGTCCTGCA | ATATGATGAA | CAGGAATAAG | 1320 |
| | TGGCTnATCG | TAAGCAAATG | CCAATGCTTT | GGCTGCATTA | ACACCTATTA | GTAACGCACC | 1380 |
| | AATTAGTCCA | GGGCCTTCTG | TAACCGCTAT | GGCATCAATA | TCTTCTATTG | ATACATCGGC | 1440 |
| 20 | ATCCCCTAGA | GCCTCGTTTA | TTGTTGCTGT | TATACCTTCA | ACGTGATGTC | TACTTGCCAC | 1500 |
| | TTCGGGAACG | ACACCGCCAA | ATCGTTTATG | ACTTTCAATC | TGACTTAAAA | CTGTATTTGA | 1560 |
| | таааататст | CTGCCATTTT | TTATAACACT | AACGCTTGTT | TCATCACAAC | TTGTTTCAAC | 1620 |
| ?5 | AGCTAGTATT | AATATATCTT | TAGTCATTTA | AATTCACCCA | CATAACCATT | GCGTCCTCAC | 1680 |
| | CTTCACCATA | ATAATTTTTA | CGTTTACCAC | CATATTGAAA | TCCTAAATTT | TCATATACAT | 1740 |
| | GTTGTGCCAC | TTTATTATTA | ACTCTTACTT | CTAAACTCAT | CACATCACAA | GTGTGACTTG | 1800 |
| 30 | CATAGTTTAT | TCCGTATTTT | AAAAGCATTT | GACCTAAACC | ATAGCCTCTA | TAATTATCAT | 1860 |
| | CGATTGCAAC | TGTTGTAATT | TGAGCTTGAT | CGATAACAAT | CCATAAACCT | AAATAACCAA | 1920 |
| 15 | TAATTTGTTG | TTCAAATTCt | AAGACAAAAT | ATTTCGCAAA | GTTATTTTGC | TCTATTTCAT | 1980 |
| | GATAAAATGC | GTCAATTGTC | CAAGAACTGT | CATTGAAACT | CCGACGCTCA | AGATCAAAGA | 2040 |
| | CTTGTGGCAC | ATCTTCTTTA | GTCATCTCTC | TAATGTTTAA | TTGTTCTTTT | GACTGTTGAT | 2100 |
| | CCAATITCGT | TCCGCCTCAG | CTAATTTATG | GTATTTAGGA | GTAAATGTAT | GTACGTCTGA | 2160 |
| | AGGTTTATCT | AGCAATTGAT | ACATGACTGA | TGCATTTGGT | AGCTGCGCAA | TCACTTCACC | 2220 |
| | TTGTAATTCA | TCTTGTAATT | TTACAGTATC | TTTCCCAATA | TAAATAAATG | GTTGGTTTAA | 2280 |
| 15 | ATCTTCTAAA | AAAGCTCGCA | ATGCCTCTAT | CGACATATAT | TGATCTTCTA | AAATAGTCAC | 2340 |
| | TAATTGACCA | TTTTGCCACT | GGAATATGCC | TGTATAAACT | GCTTGTCGTC | TTGCATCAAA | 2400 |
| | CACAGGAACC | AATAATTTAT | CAGTATGATC | GATTGTTGCT | GCCAATGCCT | TTAATGATGA | 2460 |
| io | AACACCATAT | AATTTAACAT | CTAACGCATA | CGCTAATGTT | TTAGCAACAG | TAACACCGAT | 2520 |
| | ACGTAAGCCA | GTATATGAAC | CACCACCTTC | ACCARCANTA | ATCCCATCTA | Names Consorted | 2590 |

| | TTGTTTAGAA | TCCGTAGTTA | TTTCAGCTAA | AACTTCATCG | TTTTGCATCA | ATGCTACTGA | 2700 |
|----|--------------|------------|------------|------------|------------|------------|--------|
| | TAATGGTTGA | TTCGATGTAT | CAATGAGCAG | CGAATTCATG | GATAATTGCC | TCCTTAATTT | 2760 |
| 5 | GTTCATAATG | TTCTCCTTGC | GCGAACAACT | CAATTTGTCT | TGTATTTTCA | GATATTGTTG | . 2820 |
| | AAATGTTAAT | AGATAAATGC | GTCGCTGGAA | GTAAATCTTT | TATAAATTGA | CTCCATTCAA | 2880 |
| 10 | TAACAGTAAT | TGCCTGATCT | TCGAAAAATT | CATCAAATCC | TAAATCTTCA | TCAGAATCTT | 2940 |
| 70 | CTAAGCGATA | ACAATCCATA | TGATGCAATT | TTAAATTTTT | ACCCCTATAT | GATTTAATGA | 3000 |
| | TGTTAAATGT | CGGGGAATTA | ATCGTACGTC | TTACACCAAG | AGCTTTTCCT | ATAAATTGCG | 3060 |
| 15 | TTAACGTTGT | TTTACCTGCT | CCTAAATCTC | CGTTAAGTAA | AATCAAATCA | CCACTTTTCA | 3120 |
| | ATTGCTCAAC | TAAAAATATA | GCAAATTGAT | TCATTTCATC | TAAATTATTT | ATCTTTATCA | 3180 |
| | ATGTTGATTC | TCCTATATTA | TGCTTTTCAT | TCATAAAAAT | GATTATCCAT | TGTTCAATCG | 3240 |
| 20 | TATCTAACTT | TATATTTAAC | CTTTATATTG | TAACAAATTT | CAACTTAAAT | TTCTTATCTT | 3300 |
| | TGAAACAGAT | TATCTATTCA | AAGTTAATTG | TAAGAAAATT | TAAAATATTT | GTTGACATAC | 3360 |
| | TAAAGCAGAT | ATAGTAAATT | AAATTTATCA | AATTTTTAGA | CAATTCTAAC | TATTAAAGTG | 3420 |
| 25 | ATATATACCA | TTCACGGAAG | GAGTATAATA | AAATGCTTAA | TCAATATACT | GAACATCAAC | 3480 |
| | CGACAACTTC | AAATATTATT | ATTTTATTAT | ACTCTTTAGG | ACTCGAACGT | TAGTAAATAT | 3540 |
| | TTACTAAACG | CTTTAAGTCC | TATTTCTGTT | TGAATGGGAC | TTGTAAACGT | CCCAATAATA | 3600 |
| 30 | TTGGGACGTT | TTTTTATGTT | TTATCTTTCA | ATTACTTATT | TTTATTACTA | TAAAACATGA | 3660 |
| | TTAATCATTA | AAATTTACGG | GGGAATTTAC | TATGCGAaCG | AgcATGATCA | AAAAAGGAGA | 3720 |
| 35 | TCACCAAGCA | CCAGCAAGAA | GTCTTTTACA | TGCCACGGGC | GCGCTAAAAA | GTCCAACTGA | 3780 |
| | TATGAACAAA | CCATTTGTAG | CTATTTGTAA | CTCTTATATT | GATATTGTTC | CTGGACATGT | 3840 |
| | TCACTTGAGA | GAGCTTGCAG | ATATAGCTAA | AGAAGCAATT | AGAGAAGCCG | GTGCCATTCC | 3900 |
| 40 | ATTTGAATTC | AATACAATTG | GTGTTGATGA | TGGAATAGCT | ATGGGACATA | TCGGAATGCG | 3960 |
| | ATATTCTCTA | CCATCACGTG | AAATTATTGC | AGATGCAGCT | GAAACTGTAA | TTAACGCTCA | 4020 |
| | TTGGTTTGAC | GGCGTATTTT | ACATTCCTAA | TTGTGACAAG | ATTACACCCG | GTATGATTTT | 4080 |
| 45 | AGCAGCCATG . | AGGACAAACG | TACCAGCTAT | CTTTTGCTCT | GGTGGACCAA | TGAAAGCTGG | 4140 |
| | | | | • | | TCGGCGCATT | |
| | | | | | | CCTGCCCTAC | 4260 |
| 50 | | | * | | | TGGAAGTTTT | |
| | AGGTCTAGCA | | | | | | |

| | TATCGTTACT | CGCGAAgCAA | TTGATGATGC | ATTTGCACTT | GATATGGCTA | TGGGTGGTTC | 4500 |
|-----|------------|------------|------------|------------|------------|--------------|------|
| | AACAAACACG | GTACTGCATA | CGTTAGCCAT | TGCCAATGAA | GCTGGTATTG | ATTATGACTT | 4560 |
| 5 | AGAGCGCATT | AATGCTATTG | CCAAACGCAC | GCCATATTTA | TCAAAAATAG | CACCTAGTTC | 4620 |
| | ATCGTATTCA | ATGCATGATG | TGCATGAAGC | TGGTGGCGTC | CCAGCAATTA | TTAATGAATT | 4680 |
| 10 | GATGAAGAAA | GATGGCACGT | TACACCCAGA | TAGAATCACA | GTTACTGGCA | AAACGTTACG | 4740 |
| 70 | TGAAAATAAC | GAAGGCAAAG | AAATTAAGAA | CTTTGATGTC | ATTCACCCTC | TTGATGCACC . | 4800 |
| | ATATGATGCA | CAAGGCGGTT | TATCTATCTT | ATTTGGTAAT | ATCGCCCCTA | AAGGCGCAGT | 4860 |
| 15 | TATTAAAGTT | GCCGCCTTG | ATCCATCTAT | CAAAACATTT | ACTGGGAAAG | CAATTTGTTT | 4920 |
| | CAATTCGCAT | GATGAAGCTG | TTGAAGCAAT | AGACAATCGT | ACCGTTCGTG | CAGGCCACGT | 4980 |
| | CGTTGTCATT | AGATATGAAG | GACCTAAAGG | TGGACCAGGT | ATGCCTGAAA | TGTTAGCACC | 5040 |
| 20 | TACTTCCTCT | ATTGTTGGTC | GCGGCTTAGG | TAAAGATGTT | GCATTAATTA | CTGATGGGCG | 5100 |
| | TTTTTCCGGT | GCCACAAGAG | GTATTGCAGT | TGGTCATATT | TCCCCTGAAG | CTGCATCTGG | 5160 |
| | TGGACCAATT | GCCTTAATTG | AAGATGGTGA | TGAGATTACT | ATTGATTTAA | CAAATCGTAC | 5220 |
| ?5 | ATTAAACGTA | AACCAGCCTG | AAGATGTTCT | AGCGCGTCGC | CGAGAATCTT | TAACACCATT | 5280 |
| | TAAAGCGAAA | GTAAAAACAG | GTTATCTAGC | TCGTTATACT | GCCCTAGTAA | CTAGCGCAAA | 5340 |
| 30 | TACAGGTGGC | GTCATGCAAG | TCCCTGAGAA | TTTAATTTAA | TTTATTTTTA | TATTGGAGAT | 5400 |
| , o | GGTTAAAATG | TCTAAAACTC | AACATGAAGT | AAACCAAAAT | ATTGACCCTT | TAAAAATGGC | 5460 |
| | TGAATCACTT | GAACCTGAAC | AACTAAATGA | AAAAACTTTA | AATGATATGC | GTTCAGGATC | 5520 |
| 35 | AGAAGTGCTA | GTAGAAGCTC | TACTTAAAGA | AAATGTGGAT | TATTTATTCG | GTTATCCTGG | 5580 |
| | TGGTGCCGTA | CTACCTTTAT | ATGACACGTT | TTATGATGGT | AAAATCAAAC | ATATTTTAGC | 5640 |
| | AAGACACGAA | CAAGGTGCTG | TTCATGCTGC | AGAAGGTTAT | GCACGTGTAT | CTGGTAAamT | 5700 |
| 10 | GCCTCGTTG | TAGTTACAAG | CGGTCCaGGT | GCAACTAATG | TAATGACAGG | TATTACGGAT | 5760 |
| | GCACATTGCG | ACTCTTTACC | TCTAGTTGTA | TTCACTGGAC | AAGTTGCTAC | ACCAGGCATT | 5820 |
| - | GGTAAAGATG | CATTCCAAGA | AGCGGATATT | CTATCTATGA | CTTCACCAAT | TACAAAACAA | 5880 |
| 15 | AATTATCAAG | TGAAACGTGT | TGAAGATATC | CCTAAAATCG | TACACGAAGC | TTTCCATGTA | 5940 |
| | GCTAATTCTG | GACGCAAAGG | TCCTGTAGTG | ATTGATTTTC | CAAAAGATAT | GGGTGTTTTA | 6000 |
| 50 | GCTACAAATG | TGGATTTATG | CGACGAAATC | AATATTCCAG | GTTATGAAGT | TGTTACAGAA | 6060 |
| ,u | CCAGAAAATA | AAGACATTGA | CACTTTCATC | TCACTTTTAA | AAGAAGCGAA | AAAGCCTGTC | 6120 |
| | GTATTAGCCG | GCGCAGGTAT | TAATCAATCA | AAATCAAATC | AATTATTAAC | ACAGTTTGTT | 6180 |

| | GATACACTAT | TTTTAGGTAT | GGGAGGAATG | CATGGTTCT | r atgctagtaa | CATGGCATTA | 6300 |
|-----------|---------------|------------|------------|------------|--------------|------------|------|
| | ACTGAGTGTG . | ATTTACTCAT | TAATTTAGGT | AGCCGCTTCC | ATGATAGATT | AGCAAGCAAA | 6360 |
| 5 | CCTGATGCCT | | | • | | | 6420 |
| | AATAAAGTTA ' | | | | | _ | 6480 |
| | TTAAATGATA | | | | • • | | 6540 |
| 10 | AATAAGCAGA I | | | | | | 6600 |
| | CAAACAATCG A | | | | | | 6660 |
| 15 | GGACAACATC A | * * | | | | | 6720 |
| | ACAAGCGGTG (| | | | | | 6780 |
| • | GCTAATCCTG A | | | | | | 6840 |
| 20 . | CAAGAAATGG (| | | | | | 6900 |
| | GGAACATTAG G | | | | | • | 6960 |
| | TCAGTATTTA A | | | • | | | 7020 |
| 25 | TTCTTAATCG A | | | | | • | 7080 |
| | GGACCAGCTT T | - | | | | | 7140 |
| | AGTGGCAAAT C | | | | | | 7200 |
| 30 | TTGCGGATCA A | | | | | | 7260 |
| | ATATCGATAC A | | - | • | | | 7320 |
| | AAGTCGATAT T | | | | | | 7380 |
| 35 | TTAATGTTTT A | | | | | | 7440 |
| | AAGGČAATGA T | | | | | | 7500 |
| 40 | CTATTACAAA G | | | | | | 7560 |
| | TTTATTATGA C | | | | | | 7620 |
| | AAAATTGCAG T | | | | | | 7680 |
| 45 | AATGGATATG A | • | | | | | 7740 |
| | GATGGATTTG A | | | | | | 7800 |
| | CTATTACCTG A | | | | | | 7860 |
| 50 | AAACATAATG CO | | | | | | 7920 |
| | | | | | • | AGTTAGACGT | 7980 |
| | | | | | | | |

| | CAAGCACGTA | ATATTGCTTT | AAGTTATGCA | AAAGGTATTG | GTGCAaCTCG | TGCAGGTGTT | 8100 |
|----|------------|------------|------------|------------|------------|------------|------|
| | ATTGAAACAA | CATTTAAAGA | AGAAACTGAG | ACAGATTTAT | TTGGTGAACA | AGCAGTACTT | 8160 |
| 5 | TGCGGTGGTG | TATCGAAATT | AATTCAAAGT | GGCTTTGAAA | CATTAGTAGA | AGCGGGTTAT | 8220 |
| | CAACCAGAAT | TAGCTTATTT | TGAAGTATTA | CATGAAATGA | AATTAATCGT | TGATTTGATG | 8280 |
| _ | TATGAAGGCG | GTATGGAAAA | TGTACGTTAC | TCAATTTCAA | ATACTGCTGA | ATTTGGTGAC | 8340 |
| 10 | TATGTTTCAG | GACCACGTGT | TATCACACCA | GATGTTAAAG | AAAATATGAA | AGCTGTATTA | 8400 |
| ` | ACTGATATCC | AAAATGGTAA | CTTCAGTAAT | CGCTTTATCG | AAGACAATAA | AAATGGATTC | 8460 |
| 15 | AAAGAATTTT | ATAAATTACG | CGAAGAACAA | CATGGTCATC | AAATTGAAAA | AGTTGGTCGT | 8520 |
| | GAATTACGCG | AAATGATGCC | TTTTATTAAA | TCTAAAAGCA | TTGAAAAATA | AGATAGACCT | 8580 |
| | ACAATGAGGA | GTTGTTAAAT | ATGAGTAGTC | ATATTCAAAT | TTTTGATACG | ACACTAAGAG | 8640 |
| 20 | ACGGTGaACA | AACACCAGGA | GTGAATTTTA | CTTTTGATGA | ACGCTTGCGT | ATTGCATTGC | 8700 |
| | AATTAGAAAA | ATGGGGTGTA | GATGTTATTG | AAGCTGGATT | TCCTGCTTCA | AGTACAGGTA | 8760 |
| | GCTTTAAATC | TGTTCAAGCA | ATTGCACAAA | CATTAACAAC | AACGGCTGTA | TGTGGTTTAG | 8820 |
| 25 | CTAGATGTAA | AAAATCTGAC | ATCGATGCTG | TATATGAAGC | AACAAAAGAT | GCAGCGAAgC | 8880 |
| | CGGTCGTGCA | TGTTTTTATA | GCAACATCAC | CTATTCATCT | TGAACATAAA | CTTAAAATGT | 8940 |
| | CTCAAGAAGA | CGTTTTAGCA | TCTATTAAAG | AACATGTCAC | ATACGCGAAA | CAATTATTTG | 9000 |
| 30 | ACGTTGTTCA | ATTTTCACCT | GAAGATGCAA | CGCGTACTGA | ATTACCATTC | TTAGTGAAAT | 9060 |
| | GTGTACAAAC | TGCCGTTGAC | GCTGGAGCTA | CAGTTATTAA | TATTCCTGAT | ACAGTCGGCT | 9120 |
| 35 | ACAGTTACCA | TGATGAATAT | GCACATATTT | TCAAAACCTT | AACAGAATCT | GTAACATCTT | 9180 |
| | CAAATGAAAT | TATTTATAGT | GCTCATTGCC | ATGACGATTT | AGGAATGGCT | GTTTCAAATA | 9240 |
| | GTTTÄGCTGC | AATTGAAGGC | GGTGCGAGAC | GAATTGAAGG | CACTGTAAAT | GGTATTGGTG | 9300 |
| 40 | AACGAGCAGG | TAATGCAGCA | CTTGAAGAAG | TCGCGCTTGC | ACTATACGTT | CGAAATGATC | 9360 |
| • | ATTATGGTGC | TCAAACTGCT | CTTAATCTCG | AAGAAACTAA | AAAAACATCG | GATTTAATTT | 9420 |
| | CAAGATATGC | AGGTATTCGA | GTGCCTAGAA | ATAAAGCAAT | TGTTGGCCAA | AATGCATTTA | 9480 |
| 45 | GTCATGAATC | AGGTATTCAC | CAAGATGGCG | TATTAAAACA | TCGTGAAACA | TATGAAATTA | 9540 |
| | TGACACCTCA | ACTTGTTGGT | GTAAGCACGA | CTGAACTTCC | ATTAGGAAAA | TTATCTGGTA | 9600 |
| | AACACGCCTT | CTCAGAGAAG | TTAAAAGCAT | TAGGTTATGA | CATTGATAAA | GAAGCGCAAA | 9660 |
| 50 | TAGATTTATT | TAAACAATTC | AAGGCCATTG | CGGACAAAAA | GAAATCTGTT | TCAGATAGAG | 9720 |
| | ATATTCATGC | GATTATTCAA | GGTTCTGAGC | ATGAGCATCA | AGCACTTTAT | AAATTGGAAA | 9780 |

| | AAGAGGGTCA TATTTACCAG GATTCAAGTA TTGGTACTGG TTCAATCGTA GCAATTTACA | 9900 |
|----------|---|-------|
| 5 | ATGCAGTTGA TCGTATTTTC CAGAAAGAAA CAGAATTAAT TGATTATCGT ATTAATTCTG | 9960 |
| | TCACTGAAGG TACTGATGCC CAAGCAGAAG TACATGTAAA TTTATTGATT GAAGGTAAGA | 10020 |
| | CTGTCAATGG CTTTGGTATT GATCATGATA TTTTACAAGC CTCTTGTAAA GCATACGTAG | 10080 |
| 10 | AAGCACATGC TAAATTTGCA GCTGAAAATG TTGAGAAGGT AGGTAATTAA TTATGACTTA | 10140 |
| | TAACATTGTT GCCCTACCTG GTGATGGAAT CGGTCCAGAA ATTTTGAACG GATCTCTATC | 10200 |
| | ATTGCTTGAA ATTATAAGTA ATAAATATAA CTTTAATTAT CAAATAGAGC ACCACGAATT | 10260 |
| 15 | TGGTGGTGCC TCTATTGATA CATTCGGCGA GCCTTTAACT GAGAAAACCT TAAATGCGTG | 10320 |
| | TAAAAGAGCA GATGCTATTT TACTGGGTGC AATCGGTGGA CCTAAATGGA CAGATCCTAA | 10380 |
| | CAATCGACCA GAACAAGGAT TATTAAAATT GCGTAAATCC TTAAATTTAT TTGTAAATAT | 10440 |
| 20 | ACGCCCCACT ACCGTTGTCA AAGGCGCTAG TTCTTTATCA CCTTTAAAGG AAGAACGCGT | 10500 |
| | TGAAGGCACA GATTTAGTTA TAGTCCGTGA ATTGACAAGT GGTATTTATT TTGGAGAACC | 10560 |
| | TAGACATTTT AATAATCACG AGGCCTTAGA TTCTCTTACT TATACAAGAG AAGAAATAGA | 10620 |
| ?5 | ACGCATTGTT CACGTAGCAT TTAAATTGGC CGCTTCAAGA CGAGGAAAAC TAACATCAGT | 10680 |
| | TGATAAAGAA AATGTATTAG CTTCTAGTAA ATTGTGGCGC AAAGTCGTAA ATGAAGTAAG | 10740 |
| 20 | TCAATTATAT CCAGAAGTAA CAGTAAATCA CTTATTTGTT GATGCTTGTA GTATGCATTT | 10800 |
| 30 | AATCACAAAT CCAAAACAAT TTGACGTCAT CGTATGTGAA AACTTATTTG GCGATATTTT | 10860 |
| | AAGTGATGAA GCTTCAGTGA TTCCTGGTTC ACTTGGTTTA TCACCTTCTG CTAGTTTTAG | 10920 |
| 15 | TAACGATGGT CCAAGATTGT ATGAGCCTAT TCATGGATCA GCACCAGATA TTGCAGGTAA | 10980 |
| | AAACGTTGCC AATCCATTTG GAATGATTCT ATCTTTAGCG ATGTGTTTAC GTGAAAGCTT | 11040 |
| | AAATCAACCA GATGCTGCAG ATGAATTAGA ACAACATATT TATAGCATGA TTGAACATGG | 11100 |
| o | GCAAACGACA GCAGATTTAG GCGGCAAATT GAATACTACT GATATTTTCG AAATTCTATC | 11160 |
| | TCAAAAATTG AATCACTAAG GGGGAGATGT AAATGGGTCA AACATTATTT GACAAGGTGT | 11220 |
| | GGAACAGACA TGTGTTATAC GGGAAATTGG GCGAACCGCA ACTATTATAC ATTGATTTAC | 11280 |
| 5 | ACCTTATACA TGAAGTTACT TCTCCTCAAG CATTTGAAGG ACTTAGGCTT CAAAACAGAA | 11340 |
| | AATTAAGACG CCCAGATTTA ACATTTGCAA CACTCGATCA CAATGTTCCT ACTATTGATA | 11400 |
| • | TATTCAATAT TAAAGATGAA ATTGCAAACA AACAATCAC AACATTACAA AAAAACGCCA | 11460 |
| 0 | TAGATTTTGG GGTGCATATT TTTGATATGG GTTCTGATGA ACAAGGTATT GTTCACATGG | 11520 |
| | TAGGACCTGA GACAGGACTT ACACAGCCTG GCAAGACAAT CGTTTGTGGT GACTCTCACA | 11580 |

| | ATGTTTTCGC | AACTCAAACG | CTATGGCAAA | CAAAACCCAA | AAACTTAAAA | ATCGATATTA | 11700 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | ATGGTACCTT | ACCAACAGGC | GTCTATGCTA | AGGACATTAT | TCTGCATTTA | ATTAAAACGT | 11760 |
| 5 | ATGGTGTTGA | CTTTGGTACA | GGCTATGCTT | TGGAATTTAC | TGGCGAAACA | ATTAAAAACC | 11820 |
| | TTTCAATGGA | TGGTCGAATG | ACTATTTGTA | ACATGGCTAT | CGAAGGTGGT | GCCAAATACG | 11880 |
| | GCATAATCCA | ACCTGATGAT | ATAACATTTG | AATATGTTAA | AGGGAGACCA | TTTGCCGATA | 11940 |
| 10 | ACTECGCTAA | ATCAGTTGAT | AAGTGGCGTG | Agctatattc | TGATGACGAC | GCGATATTTG | 12000 |
| | ATCGTGTAAT | TGAACTTGAT | GTTTCAACAT | TAGAACCACA | AGTGACATGG | GGAACTAATC | 12060 |
| 15 | CTGAAATGGG | TGTTAATTTC | AGTGAACCAT | TCCCTGAAAT | CAATGATATC | AACGATCAAC | 12120 |
| 13 | GTGCGTATGA | TTATATGGGG | TTAGAACCAG | GTCAAAAAGC | TGAAGACATC | GACTTAGGGT | 12180 |
| | ATGTTTTTCT | CGGTTCATGT | ACAAATGCTA | GACTATCAGA | TTTGATTGAA | GCTAGTCATA | 12240 |
| 20 | TTGTTAAAGG | AAATAAAGTT | CATCCAAATA | TTACAGCTAT | TGTCGTACCA | GGTTCTCGTA | 12300 |
| | CAGTAAAAAA | AGAAGCAGAA | AAATTAGGTC | TAGATACTAT | CTTTAAAAAT | GCAGGATTTG | 12360 |
| | AATGGCGTGA | ACCAGGATGT | TCAATGTGTT | TAGGCATGAA | TCCTGACCAA | GTACCTGAGG | 12420 |
| ?5 | GCGTACATTG | TGCATCTACA | AGTAATCGAA | ACTTTGAAGG | ACGACAAGGC | AAAGGTGCAA | 12480 |
| | GAACACATTT | AGTATCCCCT | GCTATGGCAG | CAGCAGCAGC | TATTCATGGT | AAATTTGTGG | 12540 |
| | ACGTAAGAAA | GGTGGTTGTT | TAAATGGCAG | CAATCAAACC | TATTACAACA | TATAAAGGTA | 12600 |
| 30 | AAATAGTCCC | TCTCTTCAAC | GACAATATCG | ATACAGACCA | AATCATTCCT | AAGGTACACT | 12660 |
| | TAAAGCGTAT | TTCAAAAAGT | GGCTTTGGTC | CATTTGCTTT | TGATGAATGG | CGGTACTTAC | 12720 |
| | CTGATGGTTC | AGATAATCCT | GATTTCAATC | CTAACAAACC | ACAATATAAA | GGGGCTTCTA | 12780 |
| 35 | TTTTAATTAC | TGGAGATAAT | TTTGGATGTG | GTTCAAGTCG | TGAACATGCT | GCTTGGGCTC | 12840 |
| | TTAAGGACTA | TGGTTTTCAT | ATTATTATTG | CAGGAAGTTT | CAGTGACATA | TTTTATATGA | 12900 |
| 40 | ATTGCACTAA | AAATGCGATG | TTGCCTATCG | TTTTAGAAAA | AAGTGCCCGT | GAACATCTTG | 12960 |
| | CACAATATGT | TGAAATTGAG | GTCGATTTAC | CAAATCAAAC | TGTGTCATCA | CCAGACAAGC | 13020 |
| | GTTTCCATTT | TGAAATTGAT | GAAACTTGGA | AGAATAAACT | TGTAAATGGC | TTAGATGACA | 13080 |
| 45 | TTGCAATCAC | CCTACAATAT | GAATCATTAA | TAGAAAAATA | TGAAAAATCa | CTTTAAGGGA | 13140 |
| | GTTGAATATT | ATGACAGTCA | AAACAACAGT | TTCTACGAAA | GATATCGATG | AGGCATTTTT | 13200 |
| | AAGACTȚAAA | GATATTGTCA | AAGAAACACC | TTTACAATTA | GACCATTACT | TATCTCAAAA | 13260 |
| 50 | GTATGATTGT | AAAGTCTATT | TAAAACGAGA | AGATTTACAA | TGGGTACGTT | CTTTTAAATT | 13320 |
| | AAGAGGTGCT | TACAACGCTA | TTTCTGTTTT | ATCAGATGAA | GCTAAAAGTA | AAGGTATTAC | 13380 |

| | AAACGCIGIT ATCTTTATGC CAGTCACTAC ACCTTTACAA AAGGTAAATC AAGTAAAGTT | - 1350 |
|----|---|-----------|
| | CTTTGGAAAT AGTAACGTTG AAGTTGTACT CACTGGTGAT ACATTTGATC ACTGTTTAGC | 1356 |
| 5 | TGAAGCTTTA ACTTATACAA GTGAACATCA AATGAACTTT ATAGATCCAT TCAATAATGT | 1362 |
| | TCATACAATT TCTGGACAAG GTACGCTTGC TAAAGAAATG CTAGAACAAG CAAAGTCTGA | 13680 |
| | CAATGTTAAC TTTGATTATC TATTTGCCGC AATTGGTGGT GGCGGTTTAA TTTCAGGTAT | 1374(|
| 10 | TAGTACTTAC TTTAAAACCT ATTCACCTAC CACGAAAATT ATAGGTGTTG AACCTTCAGG | 13800 |
| | TGCAAGTAGT ATGTATGAAT CTGTTGTGGT AAATAATCAG GTAGTCACAT TGCCTAATAT | 13860 |
| 15 | CGATAAATTT GTGGACGGTG CATCTGTAGC TAGAGTTGGC GATATTACAT TTGAAATTGC | 13920 |
| | AAAAGAAAAT GTAGATGATT ACGTTCAAGT AGATGAAGGT GCAGTTTGTT CTACGATTTT | 13980 |
| | AGATATGTAT TCAAAACAAG CAATTGTAGC AGAACCTGCT GGCGCATTAA GTGTAAGTGC | 14040 |
| 20 | GCTTGAAAAC TATAAAGATC ATATTAAAGG TAAAACAGTG GTTTGTGTCA TTAGTGGTGG | |
| | TAATAATGAT ATTAATCGAA TGAAAGAAAT TGAAGAACGT TCATTACTAT ACGAAGAAAT | 14160 |
| | GAAGCATTAC TITATCTTAA ATTTCCCTCA ACGTCCAGGT GCATTGAGAG AATTTGTAAA | |
| 25 | TGACGTATTA GGACCTCAAG ACGATATTAC TAAATTTGAA TACTTAAAAA AATCTTCTCA | 14220 |
| | AAATACAGGT ACTGTCATTA TTGGTATTCA ACTTAAAGAT CATGATGATT TAATACAACT | 14340 |
| | CAAACAACGT GTAAAECATT TCGATCCTTC CAATATTTAT ATTAATGAAA ATAAGATGTT | 14400 |
| 30 | ATATTCATTG TTAATTTAAC ACATAGTAAG AAAAACAGTC ATAAATTGAT TTCTAATTGA | 14460 |
| | AATCATCTTA TGACTGCTTT TTATTATACT TTACATTTCT CGTTTCGTCA GATTCAAACG | 14520 |
| | TTTTCACTTC GCCAAGCCAT CTTTCTTTGT GTTTGCTTTT aTTTTGACGT TTTAGACATA | 14580 |
| 35 | AAAAAAGAGA CCTTGCGGTC TCAATGCGGC TCATCGCATC CACTTTTTGC CTGGCAACGT | |
| | TCTACTCTAG CGGAACGTAA GTTCGACTAC CATCGACGCT AAGGAGCTTA ACTTCTGTGT | 14640 |
| O | TCGGCATGGG AACAGGTGTG ACCTCCTTGC TATAGTCACC AGACATATGA ATGTAATTTA | 14700 |
| | TACATTCAAA ACTAGATAGT AAGTAAAAGT GATTTTGCTT CGCAAAACAT TTATTTTGAT | |
| | TAAGTCTTCG ATCGATTAGT ATTCGTCAGC TCCACATGTC ACCATGCTTC CACCTCGAAC | |
| 5 | CTATTAACCT CATCATCTTT GAGGGATCTT ATAACCGAAG TTGGGAAATC TCATCTTGAG | 14880 |
| | GGGGGCTTCA TGCTTAGATG CTTTCAGCAC TTATCCCGTC CACACATAGC TACCCAGCTA | 14940 |
| | TGCCGTTGGC ACGACAACTG GTACACCAGA GGTATGTCCA TCCCGGTCCT CTCGTACTAA | 15000 |
| 0 | GGACAGCTCC TCTCAAATTT CCTACGCCCA CGACGGATAG GGACCGAACT GTCTCACGAC | 15060 |
| | GTTCTGAACC CAGCTCGCGT ACCGCTTTAA TGGGCGAACA GCCCAACCCT TGGGACCGAC | 15120 |
| | Section occurred indeaction | 15180 |

| | GTGGAACTT | 15249 |
|----|---|-------|
| | (2) INFORMATION FOR SEQ ID NO: 103: | |
| 5 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 14051 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 10 | (b) Topologi: Timear | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 103: | |
| | | |
| 15 | GTGGCAATAT TTCTAGTTCT CGTTTTGATA AGATTTTAAA AGGATCTGTT GTGTTTGCAG | 60 |
| | TGTCCTGATT TGAATTAGAT ACAAATTCAT TCACTAAAGA TGTTGTAAGT TTCATATCTA | 120 |
| | CATATGTTTC ACCTTTATAT ACAGTTCGAA TAGCTAACAA TAATTGTTCA TCAGGTGCAT | 180 |
| 20 | TTTTCAATAT GTAACCTTTC GCACCATTAC GCAACACATG GAACAAATAC TCCTCATCAT | 240 |
| | CAAACATTGT TAATATTAGT ATTTTAGTTT CAGGAAAACT GTCAGCAATT TTACTCGTAG | 300 |
| | CGATAAGACC TGACTCACCT GGTGGCATAC TTAAATCCAT TAGTAACACA TCAGGTTTAE | 360 |
| 25 | ATTCCATTAC TTTTTGGTAA GCTTCGACGC CATCTGCAGC CGTTGCAACA ACTTCCATAT | 420 |
| | CATTTTGATA ATTTAAAATC ATAGAGAACC CCGTACGGAC AACAGCGTGA TCATCGGCAA | 480 |
| | TGACTATTTT CAATTTTATT CCCCCAATGT ATGTTTCAAA TTGGAATGTT CAATGTAACA | 540 |
| 30 | TTGGTACCCT CACCAATTTT CGTTTCAATA TTGACGCTAC CGCTGACTAA CTCAGCTCGC | 600 |
| | TCATTCATTC CATATAAACC GAGTCCAGAA CCTTTAGGCT TAGAACTTGG ATCAAAACCA | 660 |
| | TTTCCCGCAT CTATCACTTC TGCTACCAAA TGGCGCCCAG TTTGACGGAT ACCTACATTT | 720 |
| 35 | ATTTCATTTA CATCAGCGTA TTTCAACGCA TTTAAAATAG CTTCTTGCAC TACTCGATAA | 780 |
| | ACAACCGTTT CAATATCACT ATCAAAGCGA GTATTTTTAA TATTTGATGT ATATATGATT | 840 |
| 10 | TTTATTCCAT AATTTTCTTC AAACTGTTTA AAATATGATT TAAAAGCTGC TTCAAGGCCT | 900 |
| +0 | AGATCATCCA AAGAAGCGGG TCTTAATTCA ACCGACATAT TACGTATATC ATCAATTAAT | 960 |
| | TTAGCGACAA TATATTCAAT ATTTTCTGCG TCTTCCAAAA GCTTAGTTGT ATCTTCTTGA | 1020 |
| 45 | TATTTTAATA ATCTCAATTG AACATCTACA TTGAGCATTT CTTGAATCAC ACTATCATGT | 1080 |
| | AACTCTCTAG AAATTCGCTT TCTTTCATTT TCTTGGGCTG AGATTGTTTT ACGCATCATA | |
| | | |
| 50 | CGTTGTTGAT GCAATTTCTC TTGCTGTTCA ATTTGTGATG AAACATTTTG AAGCGTAAAT | 1200 |
| | GCATGAATTC CCCTGTCTTG ATCAATCAAC TGATATGTTG CTGTAAATGG CATCACTTTT | |
| | ቸርልቸርፕፕፕሮር ፕሮፕፕሮልባልል ጥልሮሞሃርርልልል ፕፕሮርጥልጋር፣ም ርጥልሮሞምርርልም ለርልሞምርጥልልር | 1320 |

55

GTGGAACTT

| | ATCGCATTCG | CCACAGCACT | GTAATTATCT | TCTTCAGATA | ATATATCTTT | AGCAGCATCA | 1440 |
|------------|----------------|-------------|------------|-----------------------|-----------------|------------|------|
| | TTCATTGCAA | TAATTTTACC | GTTATCATCA | GCAAAAACTA | TCTTTTCGAT | TGAATGCTCA | 1500 |
| 5 | TAATATTTT | TCAATAAAGT | ATCTAACTGT | ATACTGTCCT | CATTAATCAT | GACTTACACC | 1560 |
| | CTAATTCATC | TCATTATTTA | TCATCATTGA | AAATACCAAA | CTTACGTTGA | ATATCATCAT | 1620 |
| 10 | TATCAAATAT | TTTTGGTAAA | GGACGACCAT | CTCTTTGACC | AAATAATAGT | ACGCCATACA | 1680 |
| 10 | CTTGATTCTT | ATACCAAAGC | GGCACTGCTA | AAACTGCTGT | TAATGATTCG | СТСААТАААА | 1740 |
| | TTGGATAGTC | AATCTTTTCT | TCAGGCCCTA | AAGCTAAACC | AACATTGGCT | ATTACCATAC | 1800 |
| 15 | GCTTTCCTGT | TTTCATAACA | GTTCCAGCTA | ATCCACGACC | TTTTCTTAAA | ATAATCAATT | 1860 |
| | TAAATCGATT | ATTTTTTATTA | CCTGAAACAT | AGTGCCATTT | TATTGGAGAT | GATGGTTTGT | 1920 |
| • | TAGATTCATA | GAAAGCGATT | GCCGCAAAAT | CATAACCCTC | TTCTTTGCGT | ATTTTATCTA | 1980 |
| 20 | ATGTCTCTTG | AAATCTACGA | TCTTCAATTA | TTGCTTCTGG | TGTCAAATCC | TTTCACCTCT | 2040 |
| | TATGCTTACA | CTTTATTCTT | ACGGTAAATA | ATATATCTGC | GATTTATATA | TGTCAAAGGT | 2100 |
| | ACACTCCAAA | CATGCACCAA | ACGTGTAAAT | GGCCAACAAG | `CCATAATAGT | GAAACCTAAC | 2160 |
| 25 | AATATATGCA | TTTTAAATGC | AATCGGCACA | CCACTCATCA | ATGACGCATC | TGGTTTTAAC | 2220 |
| | ATAAATAATT | GTCTAAACCA | AATTGATÄAT | GAAGTTCTGT | AGTTAAAGTC | TGGATGTTGT | 2280 |
| | ATATTTGTTA | CTAATGTTGC | GTAACATCCC | ATAAATACGA | TAAGTAATAA | TAAGAAATTT | 2340 |
| 30 | ACAAATATAT | CCGACGCTGA | ACTTAATCTT | CGAATACTTT | TCGTAGTAAC | ACGTCTCGCT | 2400 |
| | GTTAATAAAA | ACATCCCTAT | CAAAGTTATT | ATACCAAAGA | TGCTACCAAT | ATAAACAGCG | 2460 |
| 35 | CCTATATGAT | ATAAATGCTC | AGACACACCC | ACTGCATCCA | TCCATGGTTT | CGGTATTAAC | 2520 |
| | AATCCAACTA | CGTGTCCAAA | AAACACTGGA | ATAATACCTA | AGTGAAATAA | TAAACTTCCC | 2580 |
| | CACATCAACC | TTTTTCTTTC | TATTAATTCA | CTAGATTTAG | CTGTCCAAGA | AAATTTATCA | 2640 |
| 40 | TAACGATAAC | GTGCAATATG | ACCTGCGACA | AAGACAACTA | AACATAAATA | CGGAAATATA | 2700 |
| | ACCCATAAAA | ACTGATTAAG | CATGATGTTT | CACTCCTTTT | GGTGATGTCA | AACATAATTT | 2760 |
| | CAATGTTTTT | CTAAGTGCTT | GAATCACATA | GGCATATGGA | TTGTTATCTT | CACCAAGTGC | 2820 |
| 45 | ATTCGCCATC | ACATATGTTC | CATCCTCAAT | AATCATAATG | attaattgaa | TATTCTCTTC | 2880 |
| | AGCTCTTGGA | TCATTTCGCC | ATTCTGCCAC | TTGCAAAAAT | TGAAGCATCA | ACGGTAGATA | 2940 |
| | ATCAGAAAGT | TCATTATCTA | CCATTTCTAG | TCCAAACATT | TCATATAATA | CCTTTAATTT | 3000 |
| 5 0 | AGCTAACATT | TGCCCACGTT | CTTTTTGCGT | ATCAAATTTG | TTATACGTCA | TATATAATGG | 3060 |
| • | المالملململيات | CTAAAATCAA | ATCTATOTOT | 3 T 3 3 3 T C C C C T | CALCA Y CALCALO | AMA AMO | |

| | TGTTTCTTCA | AAAGTTTTTG | GATGAAAAGT | TAATTTTTCT | GGAAAACATA | ACTGTTGTGC | 324 |
|----|------------|------------|----------------|------------|---------------|--------------|------|
| | CATATATCCA | AAACTTTCTT | GATATTTTT | AAAATTATCG | AAATTAATCA | CGGAAAATCC | 3300 |
| 5 | CTCCATAGAA | ATTCTCATTA | TAAATTTCTT | GACCAGTTTT | CCCTGAACCT | ACTGCAACGC | 3360 |
| | CACAGCCTTC | ACAGTTATCT | CCAAAATGCT | CGCCGCCGTA | ATTGTATCCT | GTACTACCTT | 3420 |
| 10 | GTGCGTGATA | CGTATCTAAA | TAGGTTTCTT | TGTGTGATGT | TGGAATAACA | AATCGATCTT | 3480 |
| | CATATTTGGC | TAGTCCTAAT | AAACGATACA | TGTCTTTAGT | TTGGCGCTCG | GTTATACCTA | 3540 |
| | ATCGCTCTAA | TCGAGACGTG | TCAAATGGCT | GTTGAGTAAC | TTGAGATCTC | ATATAACTTC | 3600 |
| 15 | TCATCATTGC | CATACGTTGT | AGGGCTCCTT | TTACTGGCTC | TGTATCTCCT | GCAGTGAAAA | 3660 |
| | TATTAGCTAA | GTATTCAATA | GGTAAACGCA | TTTCTTCAAT | GGCTGGGAAA | ATCGCATCTG | 3720 |
| | GATTTTGAGT | TGTATTTTTA | CCTTCAAAAT | AGCTCATAAT | TGGGCTAAGT | GGTGGGCAAT | 3780 |
| 20 | ACCAAACCAT | CGGCATCGTT | CTAAATTCAG | GATGTAACGG | AAATGCAAGT | TTATATTCAA | 3840 |
| | TTGCTAACTT | ATAAATTGGA | GAGTTTTGTG | CAGCTTCAAT | CCAATCGTAA | CCAATACCAT | 3900 |
| | CTTTTTCAGC | TTGAGCAATG | ACTTCTTCGT | CAAATGGGTT | TAAGAATATA | TCTAATTGTT | 3960 |
| 25 | TTTCATATAA | ATCTTTCTCG | TCTACTGCTG | AAGCTGCTTC | ATGAACTCGA | TCTGCATCAT | 4020 |
| | ATAATAAAAC | ACCTAAGTAA | CGCATACGTC | CTGTACAAGT | TTCAGAGCAT | ACCGTAGGCA | 4080 |
| 30 | TACCCGCCTC | GATTCTCGGG | AAACAGAAAG | TACACTTTTC | AGCTTTGTTC | GTTTTCCAAT | 4140 |
| 30 | TGAAGTAAAC | TTTCTTATAT | GGACAACCTG | TCATACAGTA | ACGCCATCCA | CGACATGCGT | 4200 |
| | CTTGGTCAAC | TAATACAATG | CCATCTTCAT | CACGTTTATA | CATAGCACCT | GAAGGACACG | 4260 |
| 35 | ATGCAACGCA | ACTTGGATTC | AAGCAATGTT | CACATAAACG | TGGTAAATAC | ATCATAAAAG | 4320 |
| | TTTCGTCAAA | TTGGAATTTA | ATATCTTCTT | CTATTTTTTG | GATGTTAGGA | TCTTTTGGAC | 4380 |
| | CTGTAACATG | ACCACCTGCT | AAGTCATCTT | CCCAGTTAGG | TCCCCATTCA | ATTTCAATGT | 4440 |
| 10 | TATCCCCCGT | AATTTCTGAA | TACGCTCTAG | CAACTGGCGA | ATGCTTCCCT | GATTTCGCAG | 4500 |
| | TTGTTAAATG | TTCATAATTA | TAGTTCCATG | GCTCATAATA | ATCTTTAATT | AATGGCATAT | 4560 |
| | CTGGGTTATA | AAAAATTTTA | CCTAAAGCAA | TTTTTGAAAT | TCTACTTCCA | GATTTTAATT | 4620 |
| 15 | CAAGTTTCCC | TTTACGATTT | AGTACCCAAC | CACCTTTGTA | GTGTTCTTGG | TCTTCCCAAC | 4680 |
| | GTTTCGGATA | CCCTACACCT | GGCtTCGTTT | CTACGTTGTT | GAACCACATG | TACTCAGCAC | 4740 |
| | CTGGACGATT | TGTCCaAGTG | TTTTTACATG | TCACACTACA | CGTATGGCAT | CCTATGCATT | 4800 |
| 50 | TATCTAAATT | TAATACCATC | GCAACTTGCG | CTTTAATCTT | CAAGCCAATT | AACCTCCTTC | 4860 |
| | ATCTTTCTAA | СТССТАСАТА | TA A A TO COTT | TOOTTOOOR | TOTOCOTOCONON | 100 1000 110 | |

| | GGCGCGTTGT | GTGAACCACC | ACGTGTATCT | GTAATTTCTG | ACCCAGGCGT | TTGAATATGT | 5040 |
|----|--------------------|--------------|-------------|---------------|------------|------------|-------|
| | TTATCTTGTG | CATGATACAT | AAACATTGTA | CCTTTAGGCA | TACGATGCGA | AATAACTGCT | 5100 |
| 5 | CTTGCCGTTA | CAACACCATT | ACGGTTATAC | ACTTCTAGCC | AATCATTATC | TTGGATATCG | 5160 |
| | TGTTTTTCAG | CATCTTCATT | TGATATCCAA | ACCGTTGGAC | CACCTCTAAA | TAGTGTCAAC | 5220 |
| | ATATGCTTAT | TATCTTGATA | CATTGAGTGT | ATATTCCATT | TTCCATGAGG | CGTTAAATAA | 5280 |
| 10 | CGCAGTACCA | AAGCATCTGT | ACCACCTTTA | ATTTTCTTAT | CTCTATTCCC | AAATACCATT | 5340 |
| | GGCGGCAATG | TCGGTTTATA | TACTGGTAAG | CTCTCCCCAA | ATTGTTGGAA | AACTTCGTGA | 5400 |
| 15 | TCCACATAAT | AACTTTGACG | TCCTGTTAAT | GTTCTAAAAG | GTACTAGACG | TTCTATATTC | 5460 |
| ,, | GTTGTAAATG | GTGAATATCG | TCGACCTTGT | TTATTTGAAC | CTGGGAATAC | TGCTGTCGGT | 5520 |
| | ATTACTTCTC | GTGGTTGTGA | AGTTATATTT | AAAAACGAAA | TTTTCTCAGC | AGCGCGTTCG | 5580 |
| 20 | CTAGAAATAT | CTTTTAACGG | CATTCCAGTT | TGTTCTTCGA | GATCTTCATA | TGATTTTTGT | 5640 |
| | GATAATTTAC | CATTCGTAGC | AGATGAAATA | CTTAGTATTG | CATCAGCTAC | ATTACGTGCT | 5700 |
| | GTATCAATAC | GTGGACGATT | CGCTCTCACA | GAATCATCAT | TTGTATCACT | CCACGTACCT | 5760 |
| 25 | AACATACTTT | TTAATTCTTC | ATATTGTTCA | CTGACACCGA | AACTTACACC | ATGTGCTCCA | 5820 |
| | ACTITCCCTT | TTTCAAGTAC | AGGACCAAGC | GTGACATATT | TGTCGTAAAT | TTTAGTGTAG | 5880 |
| | TCGCGTTCTA | CAATTGCAAA | GTTAGGCATT | GTACGTCCAG | GTACCGCTTC | AATTTCACCC | 5940 |
| 30 | TTCGACCAAŤ | CTTTCACTAC | GCCGTATGGT | GTTGAAATTT | CTTGCTTTGT | ATCATGACTA | 6000 |
| | AGTGGA GTTG | TCACAACATC | TTTAAACGTT | CCAGGTAAAT | AGTCTTTTGC | CATTTCTGAA | 6060 |
| | AATGCTTTTG | CCAACGTTTT | ATAAATATCC | CAGTCTGAAC | GCGATTCCCA | TAACGGATCA | 6120 |
| 35 | ATGGCAGGAT | TGAAAGGATG | TACATATGGA | TGCATATCCG | TTGATGATAA | ATCATGTTTT | 6180 |
| | TCAȚACCAAG | TCGCTGCCGG | CAAAACAATG | TCAGAATATA | ACGGTGTTGC | CGTCATTCTG | 6240 |
| 40 | AAGTCTAAAG | AGACCACTAA | ATCTAACTTA | CCTGTTGTTT | CTTCACGCCA | CGTAATTTCT | 6300 |
| | TCTGGCTTTT | CATCTTCATT | TGGTGTAGCT | AATAACCCTG | ATTTTGTGCC | AAGTAAATGC | 6360 |
| | TTCATAAAGT | ATTCTTGACC | TTTTGCAGAA | CTTGAAATTA | AGTTTGAACG | CCATATAAAT | 6420 |
| 45 | AATGATTTTG | GATGATTCTT | TTTCAAATCA | GGATCTTCTA | TTGCAAATTG | TGTTTGTTTT | 6480 |
| • | GATTTCACTT | CATCAATTGC | ACGTTGCAAA | ATCGCTTCAT | TTGAATCTAT | ACCTTCATCT | 6540 |
| | TTAGCTTCTT | CTGCAAACAA | CAAACTATTT | TTATTAAATT | GTGGATATGA | TGGTAACCAA | 660,0 |
| 50 | CCAAGTCTAG | CTGCTAAAAC | ATTATAATCA | GCTGGATGTT | GATGCTTTAA | CTCCTCTGTT | 6660 |
| | 66 | CACAMMONTS > | 1.001mcm101 | MMMO & ONCORP | | TTCCTCTCTT | C326 |

| | AATGCGACAG | TACTCCATCC | TTCAATCGGA | CGACATTTTT | CTTGTCCCAC | ATAGTGAGCC | 6840 |
|----|------------|------------|------------|------------|------------|------------|------|
| | CAACCGCCAC | CATTCACACC | TTGACAGCCA | CATAACATAA | CTAAGTTTAA | GATTGAACGA | 6900 |
| 5 | TAAATCGTAT | CTGAGTTAAA | CCAATGGTTA | ATACCCGCAC | CCATGATAAT | CATTGAACGC | 6960 |
| | CCTTCAGTAT | CGATAGCGTT | TTGCGCAAAT | TCTTTCGCTA | CTTGAATGAC | AACACTTTGT | 7020 |
| 10 | TTTACGCCTG | AAATGGCTTC | TTGCCAAGCA | GGTGTATATT | TTGATTCTGC | ATCGTCGTAT | 7080 |
| | CCTTTTGATT | CTAATTTATG | ATCAAAACGA | CGCACGCCAT | ATTGACTTGC | CATTAAGTCA | 7140 |
| | AAAATTGTAG | CAATACGGAC | TTTGTCACCA | TTTGCTAAAG | TGACTTGTCG | AGTTGGAATT | 7200 |
| 15 | GGACGATTGA | ATATCCCATC | TCCATCACTA | TCAAAGTATG | GGAATTGAAT | TGTTTCTAAT | 7260 |
| | TCGTATCCAC | CTTCTGTCAT | TGATAATGTA | GGGTTAATTT | TAGAACCATC | TTCTGTTTCT | 7320 |
| | AGTTTTAAGT | TCCACTTCTT | ACCTTCTTCC | CAACGTTGAC | CCATTGTGCC | ATTAGGTACT | 7380 |
| 20 | ACTAAACTAT | CGCTGATTGC | ATCATGAATA | ACTGGCTTCC | ATTCGCCTTG | CTCTGTTGTT | 7440 |
| | TGACCTAAGT | CACTCGCTCT | TAAAAATCGA | CCCGCTTTAT | ATCCATTTTC | ATCTTCATCC | 7500 |
| | AGCATGATAA | GAAACGGCAT | ATCTGTATAT | TGTTTAGCGT | AATTTATAAA | GCGTTCATTA | 7560 |
| ?5 | GGTTGATTAA | CATAATGTTC | TTGTAAAATA | ACATGCGTCA | TTGCTTGTGC | AATTGCAGCA | 7620 |
| | TCTGAACCAG | GATTCGGTGC | TAGCCAGTTA | TCTGCAAATT | TCACATTTTC | TGCGTAATCT | 7680 |
| | GGTGCTACTG | AAATGACTTT | TGTACCTTTA | TAGCGGACTT | CAGTCATAAA | ATGTGCATCC | 7740 |
| 30 | GGAGTACGTG | TTAAAGGTAC | ATTAGAGCCC | CACATAATAA | TGTATGATGC | GTTATACCAG | 7800 |
| | TCACTTGATT | CAGGCACATC | TGTTTGCTCT | CCCCAAATTT | GTGGAGAGGC | AGGTGGTAAA | 7860 |
| 35 | TCTGCATACC | AGTCATAAAA | ACTAAGCATT | TCACCACCAA | GCAAATTGAT | GAATCGAGCA | 7920 |
| | CCTGCTGCAT | AACTAATCAT | TGACATCGCT | GGAATAGGTG | TAAATCCTGC | GATTCGATCT | 7980 |
| | GGAČCATATT | TTTTTATTGT | ATACAGTAAT | TGTGCTGCGA | TTATCTCTGT | AACGTCTTTC | 8040 |
| 10 | CAATTTGAAC | GCACGTGCCC | TCCCATACCT | CGGGCTTGCT | TATATTGTTT | GGCTTTGTCT | 8100 |
| | TCATTTTCAA | CAATAGACGC | CCATGCAGCA | ACGCGATTAC | CATTGTTTTC | TTCTAATGCT | 8160 |
| | TCAGTCCATA | AATCCCAGAG | TTTTCCACGA | ATATATGGAT | ATTTGATTCG | AAGCGGACTG | 8220 |
| 15 | TATTCATACC | AAGAGAATGA | CGCACCTCGT | GGACATCCTC | TCGGTTCATA | TTCAGGCATA | 8280 |
| | TCCGGACCAC | AACTTGGATA | GTCAGTTTGT | TGATTTTCCC | AGGTAATCAC | ACCATTTTTC | 8340 |
| | ACAAATACTT | TCCAAGAACA | TGAGCCTGTA | CAGTTAACAC | CATGTGTTGT | TCTTACTTCT | 8400 |
| | TTATCGTGGC | TCCAACGTTC | TCTGTACATT | TTTTCCCATT | CTCTACTTTT | ACTTTCTAGG | 8460 |
| , | ATCGACCAAT | TCCCATTAAA | TTTTTCTGTT | GGCTTAAAGA | AATTCAATCC | AAATTTTCCC | 8520 |

| | TAAAATGCCC AAGACTATTG CTT | TAATTAG | ATTGTACATT | TTTTCACAAA | CATAAAATAT | 864 |
|------------|----------------------------|----------|------------|------------|---------------------|-------|
| | TAGGGAATCA CCTAATTACT TAA | AGGAATTT | CCCTATCAAT | AACGGGATTI | CATTGAAATA | 870 |
| 5 | ATACACAATC ATGTATGGTC ATG | CTTATTG | CCAATCTAAA | TCGTTCAAAT | TTGGCACAAC | 876 |
| | GACAAATAAG GCTTCAACAC GAA | TATATTC | TCTCGGTTGA | AACCTTACTT | ATTCATTTAT | 882 |
| 10 | TTTTTATAAA TTAGTGACAT AAC | ACTGTAT | TAGCATCTGC | ACGATCGGTT | GAAATATATG | 888 |
| ,, | TTACATTTTC TTGCTGCTTA ATA | AATGCAT | CATAGTAATC | ATATTGCGAC | GAATGATATG | 894 |
| | TGCCATTCGA TGTATCATTT GGG | TTTAGCA | AACAGCCATA | ACCTTCGTCA | TATAAATGTT | 900 |
| 15 | CACAGAGCAT AAGGGCGTCA TGT | TTAGAAC | CACTTACTAC | ATAAAATTGC | TTCATAGGAT | 906 |
| | CATATGATTT AGGAGTGTTT TCA | GTATAAT | CAACAACTTC | CCCTATAATA | CATATACCTG | 9120 |
| | GTTTCGCCTC AATTGAATAG TGT | TGCAATT | TTGAAATAAT | ATTACTTAAA | CGCCCCTTAA | 9180 |
| 20 | CAACAAACTC GTTAAAACAC GAT | GCTTGAA | AGACAATCGC | TATCGGGTAA | TCAATATCTG | 9240 |
| | TGTATTGTTG TATCTGTGTG ATA | ATTTTCC | CTAAACGTTT | TACCCCCATA | TAAATTGCTA | 9300 |
| | ACGTGCCACC ATTCACTAAG GAA | ITGACAT | CCACTTCATT | TTCTTCTGAA | TCTTTAAAGT | 9360 |
| 25 | GACCTGTAGA AAATGTCACA CTT | TTAGCAA | CTGTACGCAT | TGTCAAACCT | GTCTGCATAG | 9420 |
| | TAGCAACTGC EGCGCTCGCT GATC | STCACCC | CTGGTACAAT | TTCAAACGCA | ATATGATGTT | 9480 |
| 3 0 | CATTTAGTAT GTCGACTTCT TCTT | rgcacac | GACCAAATAT | CGCTGGATCG | CCACCTTTAA | 9540 |
| ,,, | GTCTAACAAC CTTGTTATAT CGAC | CCCCTG | CTTCCACGAT | ACAGTCATTT | ATTITTTCTT | 9600 |
| | GCTGAATATG TTTTGCATAC GGCT | TTTTAC | CAACATCGAT | AATTTCAGTA | GTCAAATTCĠ | 9660 |
| 35 | CATATTGTAA AATTAACGGA TTCA | ACTAATC | GATCATATAG | AATGACATCC | gCTTCACGTA | 9720 |
| | TTAAACGCTC AGCCTTTTTC GTCA | TAATAAT | TCGGATTACC | TGGACCCGCA | CCTATCAAGT | 9780 |
| | AAACCTTGCC ATATTCCTCT ACAG | ACATAT | ATATACGTTC | CCGTCTGTAA | CTTCTACCTC | 9840 |
| 10 | ATAAACATCT ACACAACCTT CATC | AGGTTC 1 | TTGAACAATA | CCTGTATTTA | AÁTCAATTTT | 9900 |
| | TTGATCGTGG AGCGGGCAAA ATAC | ATATTC (| CCCACTCACT | GTCCCTTCAG | ACAATGGTCC | 9960 |
| | TTGTTTGTGT GGACAGATAT TGTG | AATCGC / | ATGAATTTTG | CCACTTTCTG | ТТАААААСАА | 10020 |
| 5 | CCCTACCTCT TTGCCTTTGA CAAT | AACCIT 1 | TTTCCAATT | AGGGGTGTTA | ATTCATCTAT | 10080 |
| | AGTTGTCACT TTAATTTTTT CTTT | TGTTTC (| CATGTATTAC | ACCTTCTCCA | CTTCAAAAAT | 10140 |
| _ | TCTACGTGCT TGAGCATTGC TAGT | TATTGC 1 | TTCCCAAGGT | TCAGCTTCGA | CTGCTTTTTT | 10200 |
| 0 | AGCATCCATA ATGCGTTCAA ATAG | TTCATT 1 | TTGTCTTTCT | GGGTCAAGTA | AGACTTCTTT | 10260 |
| | TACATTTTCA AATCCAAGTC TTCT | TAACCA 1 | GGCGCTGTT | СТТТСАССАТ | מיים ביים מיים מיים | 10330 |

| | AGTTGTTAAA | AATTCAGCTT | TTTCAACTTC | TGTACCACCA | TTACCACCGA | TATAGATTTG | 10440 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | GAATCCATTT | TCAACTGAGA | TAATACCAAA | ATCTTTAACA | CCTGATTCAA | CACAACTTCT | 10500 |
| 5 | TGGGCAGCCT | GATACACCCA | TTTTGAATTT | ATGAGGTGTA | TCGATGTATT | CAAATGTTTT | 10560 |
| | TTCTAAACGA | ATGCCAAGTC | GTGTCGTGTA | TTGCGTACCA | AATCGACAAA | ACTCTTTACC | 10620 |
| | AACACAGCTT | TTAACTGAGC | GTGTTTTCTT | ACCATAAGCT | GATGCTGAAC | GCATACCTAG | 10680 |
| 10 | GTCTTCCCAT | ATATTTGGTA | ATTCTTCTTT | TTTAACTCCA | TACAAACCAA | CACGTTGTGA | 10740 |
| | ACCTGTCACT | TTAACTAGTG | GCACATGATA | TTTCTTAGCC | ACTTCTCCTA | GACGAATCAG | 10800 |
| 15 | TTGGTCTGCA | TCTGTAACAC | CCCCACGCAT | TTGAGGTATA | ACAGAAAATG | TACCATCATT | 10860 |
| | TTGAATATTC | GCATGGTAAC | GTTCGTTAGC | AAATCTTGAT | TCTCTTTCAT | CTTCATGATC | 10920 |
| | ATGTGGATAA | ACCATGTTTA | AATAATAGTT | GATTGCTGGT | CGACATTTTG | GACATCCACC | 10980 |
| 20 | TTTATTTTTA | AAGTTTAAAA | CATGTCGAAC | TTCTTTAGAT | GTTTTTAAAC | CTTTCGCTCT | 11040 |
| | TATTTGCGTT | ACTATTTGAT | CGCGTGTCAA | ATCAGTACAA | ĆCACATATAC | CAGCAGGTTT | 11100 |
| | TGCGGCAACA | AAGTCATCTC | CTAAGGTGTG | CTGCAATATT | TGAGCAATTT | GCGGTTTACA | 11160 |
| ?5 | TTTACCACAT | GAATTCCCCG | CTTTTGTTTT | AGCCGTTACT | TCTTCAACTG | TTGTAAAGCC | 11220 |
| | ATTTTCCGTA | ATCGCATTTA | CTATAGTACC | TTTATCAACA | CCATTACAAC | CACAAATTGT | 11280 |
| | TTCATCATCA | GCCATATCAG | CAATTGATAG | CGATGCCTCT | TCTCCACCTT | TAGTAAGCAA | 11340 |
| 30 | TGATACAAGT | GTGTAATCTT | CAGTGGATTC | ACCTTTTTTC | ATCATGTTAT | AAAAGCGTGA | 11400 |
| | ACCATCATCG | ATATCACCAT | ATAGTACTGC | ACCAACTACA | TTACCGTCTT | TTAAAAAGAT | 11460 |
| 35 | TTTTTTATAG | TTATTATCAA | CACTATTAAA | TATTTCAATA | CCTTTAATTT | CTGCATTTTC | 11520 |
| | TACAATTTGA | CCAGCACTAT | ACAAGTCACA | CCCAGAAACT | TTTAATGACG | TAAATGTTGT | 11580 |
| | TGATCCCTTG | TATCCGTTCG | TTTCTTTATT | TGTTAAATGA | TCAGCTAATA | CTTTACCTTG | 11640 |
| 10 | TTCATATAGT | GGTGCAACGA | GTCCATAAAC | TTTGCCGTTA | TGTTCTGCAC | ATTCACCAAC | 11700 |
| | TGCATATACA | TTGCTATCAC | TTGTTTGCAT | CACATCATTG | ACAACAATAC | CACGATTAAC | 11760 |
| | ATCTAGACCT | GATTCTTTGG | CTACTTCTGT | GTATGGTCGT | ATACCTACTG | CCATAACAAC | 11820 |
| 15 | TAAGTCTGCC | GGAATCTCGC | GTCCATCAGC | CAATTTAACA | CCCTCAACAT | CATCTTCTCC | 11880 |
| | TAAGATTTCA | GTTGTGTTGG | CTTGCATTTC | AAACTTCATA | CCTTGCTTTT | CTAGATCTGC | 11940 |
| | TTTAAGCATA | TTTCCAGCTT | TACGGTCTAG | TTGCATTTCC | ATCAACCATT | CAGCTAAATG | 12000 |
| 50 | TAACACCGTT | ACTTCCATAC | CTTGATCTAA | TAAACCACGT | GCACACTCTA | AACCTAGTAA | 12060 |
| | TCCTCCACCA | ATTACAATTC | | AGTCTTAGCA | ATCTTCATCA | THEFT | 12120 |

| | GAATGCTTTA GAACCTGTCG CAAAAATCAA TTTATCGTAT GATACTTCAA TACCATTTGC | - |
|--------|---|-------|
| | | 12240 |
| 5 | AGTAGTAACT GATTGATTTG CTCTATCTAC TTCAATTACA GGATCATTTG TAATTAACTC | 12300 |
| | GATACCATGT TCCTCATACC ACTCATATGG ATTCATAATT GTTTCTTCAA CTGTCATTTT | 12360 |
| | ATTTTGTAAA ATATTTGAAA GCATGATGCG GTTATAGTTT GGATAAGGTT CTTTACCTAT | 12420 |
| 10 | TACCGTAATA TCATATAAAT CGTTGGCGCG CTCTAATATT TCTTCGATTG TTCGAATGCC | 12480 |
| | CGCCATACCG TTACCAATCA TTACTAGTTT TTGCTTTGCC ATAAAATATG CCCCTTTACT | 12540 |
| | CCATAATATT TATTTCAAAA AAAGGTATTA ATTTTTCGTT AGTGCTTTTA TATTTTCATT | 12600 |
| 15 | GGAATCATTA AGCTTTCTAA TCTATCGTTA ATGATTTGCT TTAAAATTGG GTCGAAGTTA | |
| ,,, | ATTGAAGGTG TGAAGTGTAT ATCTGTATTA ATAACCATGT CATTCATTTG CTGCTTCACT | 12660 |
| • | | 12720 |
| 00 | TTGTTAACAA GTCTTCCGTC ATATAAAAAT AATGGTACGA CAATCAATTT TTGATACCGT | 12780 |
| 20 | TTCGAGATGC TTTCTAAATC ATGTGTAAAA CTAATCTCTC CATATAGCGT TCTCGCATAT | 12840 |
| | GTCGGCTTGC TAATTTGCAA ATTTTGAGCG CATATTTGTA ACTCTTCGTG TGCCTTAGTA | 12900 |
| | AACTTTCCAT TAATATTGCC GTGTGCAACA ACCATAACTC CAACTTGTTG TTCGTCACCT | 12960 |
| 25 | GCTAATGCGT CACAAATACG TTGTTCAATT AATCGTCTCA TTAAAGGATG TGTGCCAAGT | 13020 |
| | GGCTCGCTTA CTTCTACCTT TATGTCTGGA TACCGTCGTT TCATTTCATG AACGATATTC | 13080 |
| | GGTATATCCT TGAGATAATG CATTGCACTA AAGATTAGCA ATGGTACAAT TTTAAAATGG | 13140 |
| 30 | TCAACCCCAC TTTGAATCAA CGTCGTCATT ACCGTCTCTA AATCCLGATG CTCACTTTCL | 13200 |
| | AAAAACGCAA TATCATAGTG ATGTATATCA TCTTTTACTA ATTCAGAAAT AAATGCTTCT | |
| | AACGCTTGAT TCTGTCGTCC GTGCCTCATG CCATGTGCAA CAATGATATT CCCATTCACA | 13260 |
| 35 | • | 13320 |
| | TTTACCAACC CTTTCACACG TATTGTATAC CAAATCATTT TGTTTTTGTG AAAAGAATCA | 13380 |
| | CATTATAATG TAAAATCAGG GAATTCCCTG ATGCCTGTAG TCATGCATAT TCCTTATACA | 13440 |
| 10 | TTTTCCCTTT TTGTTAAATC AAAAAAAGCG ACCGATATAT GAATCCCTAC TCAACATTTA | 13500 |
| | TTTGAGCAAG CATTAATATA TCGGTCGCTT GTAGTGTATA TTATTATCTT AAAATGGTGG | 13560 |
| | TTGGCCTAAT ATTGTTTCGT CAAAGCGCTC GGGTATCAAT ACTTTGCGCA TGATCACACC | 13620 |
| 5 | TARATCGCCA TCATCATTTT CATGTTCGCT GTATATTTCA TARCCTCTTT TTTCATAAAT | 13680 |
| | TTTAAGTAAC CACGGATGCA ATCTTGCAGA TGTACCTAAA GTAACTGCCG CTGACTTTAA | 13740 |
| | CGTATCTCGC AAAAATGCTT CTTCAACATA AGTAAGTAAT TGGCTACCAT AGCCTTTCCC | |
| 0 | TTCATACTCA GGATTTGTCG CAAACCACCA GACAAAAGGA TAACCCGAAA TACTTTTCAC | |
| | | 13860 |
| | ACTTCCCCAA GGATATCTAA CCGTAATCGT AGATATAATT TCATCATCAA TTGTCATGAC | 13920 |

| CCAATCAATA CCTAGTTCTC TTAGAGGCGT AAATGCTTCA TGCATGAGTT CTTGCAATTT | 14040 |
|---|-------|
| TTCTGCATCT T | 14051 |
| (2) INFORMATION FOR SEQ ID NO: 104: | |

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1885 base pairs
 - (B) TYPE: nucleic acid(C) STRANDEDNESS: double

 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 104:

| 60 | TAGTGTATGA | CTGCTTAGTT | CGCATATGAA | ATGGCTTGGG | CTTnGATTAT | TAATCCTCAA |
|-------|------------|------------|------------|------------|------------|------------|
| 120 | CTACTGAAGA | GATAACTTTT | ACCTCGAATA | CTATCATACA | GTTCGCATGA | CATTCATACA |
| 180 | TAGCCAGACT | GTTAAACCCT | AACCGATTTT | TTCAATGGGG | TCAAGATTAC | GTTACCAATC |
| 240 | GTAAGATAAA | TGTAGATTCT | AGGTAGTCAT | AGTTTAAAGC | GGTGAAGGTG | TGCTTATAAC |
| 300 | CACCACATTT | CCTCAAAAGC | GCAAAATGTG | CAGAATACAT | AGAACACGTG | GCATTCATGT |
| 360 | AAAAATGGGC | CCTGATATCA | ATATAAACTG | CAGAACTTTT | GAAGAGATTG | GTTGAGTGAT |
| 420 | CGTATCCAGG | AATGATAAAA | AGCGAAAGAG | CGTTAGAACA | GAGAAATATG | TGATGAAGTA |
| . 480 | CAGTCCGAGA | GATACAAAAG | AGTGATAACT | GTTCAAGGAG | GTCACGGGAC | TTGGAAGCTA |
| 540 | TACTTAGCAT | GAAACCAAGT | AGATATTACA | ATAAACCTGA | GAAGCGGGTT | CAGGTTAGTT |
| 600 | AAGGCTTTAT | AAAATTGCAG | AGCATTTTCT | TCGGCAAAAA | GAAAAATTAA | TACGAATTTA |
| 660 | CAGCTATAAA | GATAAACGAC | TACCGAGTCT | TAACACTTGC | CAAGGTAAAT | AGAAAAGCCG |
| 720 | TATATAAACA | AAAAGGACGG | ATAAAAATTA | TTGACAAACT | GAAGATGATT | GCAATCTGCT |
| 780 | GCATCATATG | AAAAGTAAGA | TGATTACAGG | AAAACTAAAG | AGTATTAAAT | TGAAAGCAAA |
| 840 | TCAATCAGTT | AGCAAAGTAT | AAGGGCAAGA | AGTATGCAAG | TGAACCTCAC | CACATATTTT |
| 900 | ATAGAAGCTG | TGAACAAGCT | TAAAAGCCAT | ACAAGTACGA | Taaatcagat | TAATCATTCC |
| 960 | CTGAAACTTC | TCCTGCAAAT | GAGGCAAAGT | AGTAAGTTTG | AGGAAAAGTT | CTAAAGAAGA |
| 1020 | GCTTATTTTA | TTATCAAGAC | ATGATGTGAA | GAAAGAGAAG | TGGAGATACT | CATTACGTGA |
| 1080 | AGATTAACGG | AAACAAAATT | TTATTGACCA | GCACCTGGTA | AAGCAAACAA | TTAACGCATC |
| 1140 | TTTCCATTCA | AATCAATTTA | TTAGAGCTTC | GGTGACTATA | TATTGTAAGT | ATTCTGGAAC |
| 1200 | GTAGAAAAAG | CATTCAACTT | GATTGAACAA | ATCGCAGTTG | TAATAAGGGT | ACACAAATGG |
| 1260 | GACACTGATG | TGATGAATTA | AAGATGATTT | AGTGCAGCAG | TGGCGGTGCA | GCGAACCTCT |
| | | | | | | |

| TTGAGGTGTC | AAGAATTTGA | AATTTATGAA | TATAGATATT | GAAACATACA | GCAGTAACGA | 1380 |
|------------|------------|------------|------------|------------|------------|------|
| TATTTCGAAA | TGTGGTGCCT | ATAAATACAC | AGAAGCTGAA | GATTTCGAAA | TTTTAATTAT | 1440 |
| AGCTTATTCG | ATAGATGGTG | GAGCGATTAG | TGCGATTGAC | ATGACTAAAG | TAGATAATGA | 1500 |
| GCCTTTCCAC | GCTGATTATG | AGACGTTTAA | AATTGCTCTA | TTTGACCCTG | CTGTAAAAA | 1560 |
| GTATGCATTC | AATGCTAATT | TCGAAAGAAC | TTGTCTTGCT | AAACATTTTA | ATAAACAGAT | 1620 |
| GCCACCTGAA | GAATGGATTT | GCACAATGGT | TAATTCAATG | CGTATTGGCT | TACCTGCTTC | 1680 |
| GCTTGATAAA | GTTGGAGAAG | TTTTAAGACT | ACAAAGCCAA | AAAGATAAAG | CAGGTAAAAA | 1740 |
| TTTAATTCGT | TATTTCTCTA | TACCTTGTAA | ACCAACAAAA | GTTAATGGAG | GAAGAACTAG | 1800 |
| AAACCTACCT | GAACATGATC | TTGAAAAAtG | GCAACAATTT | ATAGATTACT | GTATTCGAGA | 1860 |
| TGTAGAAGTA | GAAATGGCGA | TTGCT | | | | 1885 |
| | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 105:

(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2656 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 105:

| 60 | ATCCATAGAA | CTCTATCTGC | CACCAGTTCC | AATTTCAAAA | TTCACTGnCA | TAATCCTTAG |
|-----|------------|------------|------------|------------|------------|------------|
| 120 | TCGATTAAGT | TGTTTGTAAC | ATTGTGATGn | ACCGGATTAT | TGTGTCAATA | ACTGNATGTT |
| 180 | CGAATAAACT | ACCTTTCCTT | TTATTCAACC | TCTACTACCA | CGAAAAATTA | TATCATCTTT |
| 240 | CAGACGGGAA | AATTGATGTG | TCTAACTAAA | TACCAAAGTT | ACKCCACCAG | CCATTTACCA |
| 300 | CTTCTAATAT | TTCGATTTTA | ACCTGGTGTA | GTGTTGTATT | CTTAATACTT | GTTATTACGT |
| 360 | ATTCTTCAGT | AATCGGTTTG | AAAATCAGTA | AGTCTTTAGG | ATACCTTTAA | CCAACCTGCA |
| 420 | TACTCTTAGG | AATTTTGTAA | TAAATCTGAT | CAACGATTTT | AAATCTAAAC | AGTGATATAG |
| 480 | TGTTTTTCTT | CATGCTCCGT | GCAGAAATTC | CGTTTTGCGG | CAATAACCGG | GATATGTTCC |
| 540 | TGTTTCTGGt | CTACCTGTGC | AAGCATATTT | GTTCGAATTT | TCAATGACAC | ATTGAAAATG |
| 600 | ACCACCCTGC | ACATCCAAGA | TCCAGGCAGT | CATTATAGTG | CTTAGAGCAC | AAGTACTTGT |
| 660 | GaCGTATAGA | TGTGTGTCTT | TAATGCATTT | ACGGGAAATC | GCTTTCGGTA | ATCTCTAAAC |
| 720 | ATTGGTAACG | TGGGTTCTGT | TTAGATATGC | GTAATATCAC | ACCAACTTCC | TATAGTAATG |
| 780 | TTCAAGAATT | TGCGATGTTT | ATACTTGAGG | TCAGTCATTG | TCCGCCTGAA | GTTTAACACG |

55

5

10

15

20

25

30

35

40

45

| | TAGTTACCCC | GATTAGAAGT | GCTTTACGTC | CTGTTTCTAG | ATCGTAATAC | ATATCTAGAC | 900 |
|-----|------------|--------------|------------|------------|------------|--------------|------|
| | CCTCAGCCTC | TTGGAAATCT | CCTTTAAAGT | TGTTATTCAC | ACCGCCTATA | TCGATGCGAC | 960 |
| 5 | GTTTAAATAA | CAATTCTTTC | GTTTTGATAT | CGAAGCCTTG | TAAGTAGTTA | GGGTTGGCTG | 1020 |
| | TATTCGAATC | ACCTGTATAC | CAATATAAGA | TACCTGCATC | ATAAGTGATA | CCTTGCATAG | 1080 |
| | GTTGTGTATC | TGAAGTGTAT | TCCATAGGTA | TATCCATTTG | ATACAATACT | TTGTCTATAC | 1140 |
| 10 | CTTTATCAAT | ATCGTCAGCA | CTTCTAACCT | CAACAAAGTT | CAACGAATTC | TTAAGTTGTC | 1200 |
| | TTTCAGTGGG | TTTATATTCA | CGTCTAAAAA | TCATTAAATT | TTCTACCGGA | TTATAAATCG | 1260 |
| 15 | CTGACGTATA | TCTGTCGTTA | AATATATTCG | GCATGACATC | TTGCATTTCA | TTACCATAAG | 1320 |
| . • | TTATTTCTCC | AGTTCTATAT | TGGAAACGTA | CAAACTTGTT | GTTTTTGTTA | CTGTCCAATA · | 1380 |
| | CAGCTGAATA | AATCCATAAT | TCTCCATCAA | TGTATCTATA | CGCATTGTGT | GTACCGTGAC | 1440 |
| 20 | CGCCGTTTTT | AACAAGCAAT | CTATCAATAA | ATTGTCCGTT | GGGCTTCAAT | CTAGATAACA | 1500 |
| | TGTAATGATT | ACCTGGACGA | GCTTGCGTCA | TATAAATAAT | TTTCGTTCTA | GGGTCTACCC | 1560 |
| | AAAATGATTG | CATTACTGCA | TTTGTATATG | GCGATAAATC | AGTGATAAAT | TCCGGTTCTT | 1620 |
| 25 | GCTCTTTTGG | TTCGAATCGG | TATTCTGTCG | CTCGATATTC | TTTATAGTGT | TCATCTACAG | 1680 |
| | CTTTCTCAAC | CTTTTTAGTG | AAAACATCTA | GTGTTGAATA | ATCATGATAC | AAACGATCTT | 1740 |
| | GCAATGTCTT | ATGACCATAA | CCTGTATTAT | CAACGCGCGC | GTCTTTTACT | TCGTTGATAC | 1800 |
| 30 | CGTCGCCGTT | ATGACCTAGT | ACCATGTTGC | TAAATCGAÇC | GTTTAAATAT | GTTAAAAAGT | 1860 |
| | CAGAGACGTT | ACTTGTAACA | TTTAAATGTT | CATACTTTAT | TTGTTCTCCA | TCATGTGCGA | 1920 |
| | ATACCTCTTT | ATTTCTGTGG | TATTCAAGAG | AGAAATTAAA | ATCCGTCAGC | ATGTCTGAAA | 1980 |
| 35 | TAAGTTTAAA | GTTATACTCA | TTTTCATCTA | CATATCTGTA | GTCAAAGACT | СТАСТТАААТ | 2040 |
| | CTGTAATTAG | TTTATTACTC | ATGTTTTCCT | CCTTTACTAT | CCATAAAACT | GATMATAATT | 2100 |
| 40 | TTTAATAAGC | TCATACATAA | TAACTTCATG | ACCTCTTTCA | TTAGGATGTA | ATCCATCAGG | 2160 |
| | CATGCTAGAT | TTTCTAAATG | CTGGATTATA | TGGTTTGAAA | TAATCTGTGT | GATAAGCATC | 2220 |
| | ATATACTGGT | ACATCCAATT | CACTACAAGC | CAATATCTGA | GCATTGACAT | AATCCTCTAA | 2280 |
| 45 | AGTTAACCCT | AGTTTGTTTT | TGTCCGTATC | TTTACGGCGT | ATCGTTGTAC | CACTCATAGG | 2340 |
| | GCATTGCCTA | GTAGCTGTCA | TTACAAGTAT | TTTTGAAGCT | GGATTATTTT | TCCTGATAAC | 2400 |
| | TTCAATTGCA | GAACAAAAGG | CGCCGTAAAA | CGTTTTAGTG | TCGGTTTTAT | CAGTGCCTAT | 2460 |
| 50 | CGGTACGCCT | GCCCAATAAC | CATGTAACCA | GTCATCATCT | GTACCTTGTA | ATATGATTAG | 2520 |
| | | 3 77770 0770 | | | | | 2500 |

| CITGCCTAAC | ATTTCT |
|------------|--------|
|------------|--------|

2656

(2) INFORMATION FOR SEQ ID NO: 106:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4854 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

10

15

20

25

30

35

40

45

50

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 106:

| AAAATGAGGG TTCTAGCGGA AATTACCAAA AGCGTGGTTC ATACTATGGG CAG | GCGTAATC 60 |
|---|--------------|
| GTATTTCAAA AGAAAAACA CCTAAATGGT TAGAAAATAG AGATAAACCT AG | TGAAGAAG 120 |
| ATTCGGCTAA AGATAATAGC GTAGATGATC AACAATTAGA GCAAGATCGA CAA | AGCATTTC 180 |
| TAGATAAATT ATCTAAAAAA TGGGAGGAGG ACAGTCAATA ATGAAGCAAT TTA | AAAGTAT 240 |
| AATTAACACG TCGCAGGACT TTGAAAAAAG AATAGAAAAG ATAAANCAGA AGT | |
| GACCCAGATG TTAAGCAATT TTTGGAAGCG CATCGAGCTG AATTMACGAA TGC | |
| GATGAAGACT TAAATGTGTT ACAAGAGTAT AAAGATCAAC AAAAACATTA TGA | |
| | |
| AAATTTGCTG ATTGTCCAAA TTTCGTAAAG GGGCATGTGC CTGAGTTATA TGT | |
| AACCGAATTA AAATACGCTA TTTACAATGC CCATGTAAAA TCAAGTACGA CGA | |
| TTTGAAGCTG AGCTAATTAC ATCTCATCAT ATGCAACGAG ATACTTTAAA TGC | CAAATTG 600 |
| AAAGATATTT ATATGAATCA TCGAGACCGT CTTGATGTAG CTATGGCAGC AGA | TGATATT 660 |
| TGTACAGCAA TAACTAATGG GGAACAAGTG AAAGGCCTTT ACCTTTATGG TCC | ATTTGGG 720 |
| ACAGGTAAAT CTTTTATTCT AGGTGCAATT GCGAATCAGC TCAAATCTAA GAA | GGTACGT 780 |
| TCGACAATTA TTTATTTACC GGAATTTATT AGAACATTAA AAGGTGGCTT TAA | AGATGGT 840 |
| TCTTTTGAAA AGAAATTACA TCGCGTAAGA GAAGCAAACA TTTTAATGCT TGA | |
| GGGGCTGAAG AAGTGACTCC ATGGGTGAGA GATGAGGTAA TTGGACCTTT GCT | |
| CGAATGGTTC ATGAATTACC AACATTCTTT AGTTCTAATT TTGACTATAG TGA | |
| | |
| CATCATTTAG CGATGACTCG TGATGGTGAA GAGAAGACTA AAGCAGCACG TAT | |
| CGTGTCAAAT CTTTGTCAAC ACCATACTTT TTATCAGGAG AAAATTTCAG AAA | |
| ATTTTAAAAT GATTGGTGTA TAATGAATAC AAATCTAAAT CGTTTAAATG ATT | |
| AGATGATCTA ATCAATATTA CACAGAAAGC CATTGTTTGA TGAGAATATG GTT | AATAAAT 1260 |
| TAGATGATTA CTACTTCATT TATGGTATTT GTAATGAATA CCCGGATCAA GACG | CGTTATC 1320 |
| | |

| | CTCGTCCCTT | GTATAGGGGC | GGGATTTTTT | GTTTTTTCA | GACATAAATG | TTTGTTGGTG | 1440 |
|----|------------|------------|------------|-----------------|-------------|------------|------|
| | TCATAAATTC | CCTGTTTATT | GTTAATAGGT | TTAATGTTAA | AACGATGATT | GTTGTTCAAT | 1500 |
| 5 | TTTTTAACGA | GGTCAGATAA | AAGTATTTAT | AAAGCAAATA | GGAGGGTTTA | ACATGGAACA | 1560 |
| | AATTAATATT | CAATTTCCAG | ATGGTAATAA | AAAGGCGTTT | GATAAAGGTA | CTACTACTGA | 1620 |
| 10 | AGATATAGCA | CAATCAATTA | GTCCTGGATT | ACGTAAAAAA | GCTGTTGCCG | GCÄAATTTAA | 1680 |
| | CGGGCAACTT | GTAGATTTAA | CTAAACCGCT | TGAAACTGAT | GGATCAATTG | AAATTGTGAC | 1740 |
| | ACCAGGTAGT | GAAGAagcGT | TAGAGGTATT | ACGTCATTCT | ACTGCACATT | TAATGGCACA | 1800 |
| 15 | CGCGATTAAA | AGGTTATATG | GTAATGTTAA | ATTTGGTGTA | GGTCCTGTAA | TAGAAGGTGG | 1860 |
| | ATTCTACTAT | GACTTCGACA | TTGACCAAAA | CATCTCATCT | GATGACTTTG | AACAAATTGA | 1920 |
| | AAAAACAATG | AAACAAATCG | TTAACGAAAA | TATGAAAATC | GAACGAAAAG | TGGTTTCACG | 1980 |
| 20 | AGATGAAGTG | AAAGAGTTAT | TCAGCAATGA | TGAATACAAA | TTAGAATTAA | TCGACGCGAT | 2040 |
| | TCCTGAAGAT | GAAAATGTAA | CATTATATAG | TCAAGGTGAT | TTTACTGATT | TATGTCGTGG | 210 |
| | AGTTCACGTT | CCATCAACAG | СТААААТТАА | AGAGTTTAAA | СТАТТАТСТА. | CAGCAGGTGC | 2160 |
| ?5 | ATACTGGCGT | GGAGATAGTA | ACAACAAAAT | GTTACAACGT | ATATACGGTA | CTGCTTTCTT | 2220 |
| | TGATAAAAA | GAATTGAAAG | CACATTTACA | AATGTTAGAA | GAGCGTAAAG | AACGTGATCA | 2280 |
| | TCGTAAAATT | GGTAAAGAGT | TAGAACTATT | CACAAATAGC | CAATTAGTTG | GTGCTGGTTT | 2340 |
| 30 | GCCATTATGG | TTACCTAACG | GTGCAACAAT | TAGACGTGAA | ATTGAACGTT | ACATTGTTGA | 2400 |
| | TAAAGAAGTT | AGCATGGGAT | ATGACCACGT | TTATACACCA | GTACTTGCTA | ATGTTGATTT | 2460 |
| 35 | ATACAAAACA | TCTGGTCACT | GGGATCACTA | TCAAGAAGAT | ATGTTCCCAC | CAATGCAGTT | 2520 |
| | AGATGAAACT | GAATCTATGG | TATTACGTCC | AATGAACTGT | CCACATCATA | TGATGATTTA | 2580 |
| | TGCGÄATAAA | CCACATTCAT | ATCGTGAATT | ACCTATCCGT | ATCGCTGAGC | TAGGAACGAT | 2640 |
| 10 | GCATAGATAT | GAAGCAAGTG | GTGCTGTATC | AGGATTACAA | CGTGTTCGTG | GTATGACTTT | 2700 |
| | AAATGATTCA | CATATCTTTG | TTCGACCTGA | TCAAATTAAA | GAAGAATTCA | AACGCGTTGT | 2760 |
| | AAACATGATT | ATTGATGTGT | ATAAAGACTT | TGGTTTCGAG | GATTATAGCT | TTAGATTAAG | 2820 |
| 15 | TTATAGAGAC | CCTGAAGATA | AAGAAAAGTA | CTTTGATGAT | GATGATATGT | GGAATAAAGC | 2880 |
| | TGAAAATATG | CTTAAAGAGG | CAGCGGATGA | GCTTGGCTTA | TCGTACGAnG | AAgCGATTGG | 2940 |
| | TGAAgCGGCA | TTCTATGGTC | CGAAACTAGA | TGTTCAAGTT | AAAACAGCGA | TGGGTAAAGA | 3000 |
| 50 | AGAGACATTA | TCAACAGCAC | AACTTGATTT | CTTATTACCA | GAACGTTTTG | ATTTAACTTA | 306 |
| | TATTCCTCAA | CATCCTCAAC | NTCNTCCTCC | A CHANCHAL PLAN | CATCOTOCTO | TTCTATCAAC | 312 |

| | AGCGCCAAAA | CAAGTTCAAA | TCATTCCAGT | TAACGTTGAT | TTACATTATO | ATTATGCGCG | 324 |
|------------|--------------|--------------|-------------|------------|------------|------------|------|
| _ | CCAATTACAA | GATGAATTGA | AATCTCAAGG | CGTTCGTGTA | AGTATTGATG | ACCGTAATGA | 330 |
| 5 | AAAAATGGGT | TATAAAATCA | GAGAAGCTCA | AATGCAAAAA | ATACCTTATO | AAATCGTAGT | 336 |
| | TGGGGATAAG | GAAGTTGAAA | ATAATCAAGT | GAATGTGCGT | CAATATGGAT | CGCAAGACCA | 342 |
| 10 | AGAAACAGTT | GAAAAAGATG | AATTTATCTG | GAATCTAGTT | GATGAAATTC | GTTTGAAAAA | 348 |
| | ACATAGATAG | ACAGTTGTCG | CAATAAAATG | CTTTAAAACT | TTTATTGCGT | ATCAAGTTTT | 3540 |
| | ACAGGGTTGA | TTATGCGTGA | TGAATCCTGT | ATATTACAAG | TTAGTTAAAA | TATTAAATTG | 3600 |
| 15 | AGTTAGAGGT | TGCATGTTTA | ATTAGTAACT | TGTCAGAAGT | ATTTATGGTA | CATAAGTTGA | 3660 |
| | ACAAGTGAAA | GGTAAAGATG | CCGAAATAGA | TATAAACCAT | AAATTATATC | TATTGGGACA | 3720 |
| | GTTTTCGAAT | AGGAACTGTA | CTGTCACAGA | ATGTGATGTG | CTACCTTATA | TAGATAATTG | 3780 |
| 20 | CCAAAGTGGT | TGCATATCTT | AAAGGTATGT | AGCCACTTTT | TTACTTTTAA | TATCACTATG | 3840 |
| | TTCTGTAAAA | AAGGGTATGA | AAGTGAATAA | AGGTTATTTA | TTTCTTGGCC | TCTAAAACAT | 3900 |
| | GGAAAGGGAG | CTTATATGTC | AAAAGTTCAA | AATGAAAGTA | ACAATGTTGT | CAAAAGGGGA | 3960 |
| 25 | CTTAAAGATC | GTCATATTTC | TATGATTGCG | ATTGGGGGTT | GTATTGGTAC | AGGTTTATTT | 4020 |
| | GTAACTTCTG | GTGGAGCAAT | TCATGATGCA | GGTGCTTTGG | GTGCATTAAT | AGGATACGCA | 4080 |
| 3 <i>0</i> | ATTATCGGAA | TAATGGTATT | TTTCTTAATG | ACGTCACTTG | GCGAAATGGC | TACGTATTTG | 4140 |
| | CCAGTATCAG | GTTCATTTAG | TACATATGCT | ACAAGATTTG | TTGATCCATC | TTTAGGGTTT | 4200 |
| | GCGCTTGGTT | GGAACTATTG | GTTTAACTGG | GTAGTGACTG | TAGCAGCAGA | TATTACGATT | 4260 |
| 35 | GCAGCACAAG | TCATTCAATA | TTGGACACCA | TTGCAAGGCA | TACCCGCTTG | GGCATGGAGT | 4320 |
| | GCGTTGTTCT | TAGTTATAAT | TTTTAGTCTG | AATTCGTTAT | CAGTTCGCGT | CTATGGTGAA | 4380 |
| | AGTGAATACT | GGTTGGCATT | GATAAAAGTG | GTTACAGTTA | TTGTTTTCAT | TGCAATTGGT | 4440 |
| 10 | TTATTAACGA | ŢTGTCGGAAT · | CATGGGTGGT | CATGTTGTAG | GATTCGAAAT | ATTTAATAAA | 4500 |
| | GGTGAAGGTC | CAATTCTTGG | TGGCAACTTA | GGAGGAAGTT | TGTTATCAAT | TCTAGGTGTA | 4560 |
| | TTCTTAATCG | CTGGTTTCTC | ATTCCAAGGT | ACTGAGTTAA | TTGGTATTAC | GGCTGGTGAA | 4620 |
| 15 | TCAGAAAATC | CTGAACGTGC ' | TGTGCCGAAA | GCAATTAAAC | AAGTATTCTG | GAGAATTTTA | 4680 |
| • | TTATTTTACA | TTTTAGCCAT (| TTTTGTTATC | GGTATGTTAA | TTCCTTATGA | TAGTAGTGCA | 4740 |
| | TTAATGGGGG (| GTAGTGATAA : | TGTAGCAACG | TCTCCATTCA | CATTAGTGTT | TAAAAATGCT | 4800 |
| i o | GGATTTGCGT | TTGCAGCATC A | ATTTATGAAT | GCAGTCATTT | TAACGTCTGT | GTTA | 4854 |
| | (2) INFORMAT | TION FOR SEC | 3 TD NO. 10 | ~ . | | | |

(A) LENGTH: 2488 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear

5 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 107:

| 10 | ATCAAAAATT | GATTGTTTTC | nATTTTTTGT | TTCAGCGCGG | GATCTTTTAC | GTCTTTTGTG | 60 |
|------------|------------|------------|--------------|------------|------------|------------|------|
| | AAAACGaTTT | TATTATTAAC | TACTTTTACT | GGATAACTTT | TGTATGTCGA | GTCAGTAGCA | 120 |
| | TTTTTTCTAT | CGTTTGTAGT | TGTGTCATAT | TCACCAGTTA | TTTTATGTGT | GTTCTTATCT | 180 |
| 15 | ACCTTTAACA | ACATACGGTC | TTCTTTTAAA | AGCTCATCTG | ATCCAACAAC | TGAATAAGAG | 240 |
| | GATTCTATAT | ACCATGTGTC | TTGATCATTA | TTTTCATAAT | GGGGATTATC | GTGACCATCA | 300 |
| | ATTTCATAAA | GCGTTTCTAA | GTTTTTAATA | GGATACGTAC | TTAGTACTTT | TTTAAGACCA | 360 |
| 20 | TCTTTCAAAT | GAATTTGTTC | CCACTTCATT | GCCAAAAACA | TATCGCCACT | GACTACAATT | 420 |
| | GAAATAATAA | TAATTGCTGC | TAAGTTTAAC | CAGAAAATTT | TATGTGCTTT | CATACATTCC | 480 |
| 25 | CACCGTTTCT | CAAAATACTT | CATTAACACT | ATAATAATAT | ATTTTGAAAA | ATATTTACAT | 540 |
| | CAGTATTAAA | GTGAATATCA | TAAATTTTAAAT | TTATGAAAAT | AATAGATATT | TATAAAAAGC | 600 |
| | GGAAAAGAGA | TACAATAAAA | AACTGCATGA | CGTTTGAGAC | GTCACACAGT | GTAACTAAAA | 660 |
| 30 | ATTTAAAAAG | TTGTTGCTAA | TTTTTCAGCA | TTATTAATAC | TAGTTGCTTT | AATTTCTTCA | 720 |
| | GTCTTATGAG | GTTCAGCATT | GTGTCCTTCA | ATAATGATTG | TTTCATATGA | TGGCACACCT | 780 |
| | AAGAATGTCA | TAATTGTTCT | TAAATAACGG | TCACCCATTT | CAAAATCAGC | AGCAGGTCCT | 840 |
| 35 | TCAGTATAAT | ATCCACCACG | TGATTGAATG | TGTAATACTT | TTTTGTCAGT | TAGTAAACCT | 900 |
| | TGTGGTCCTT | CAGCAGAATA | TTTAAAAGTT | TTACCTGCAA | TTGAAATAGC | ATCAATATAT | 960 |
| | GCTTTAACTA | CAGGTGGGAA | AGAAAGGTTC | CACATAGGCG | TTACAAATAC | ATATTTATCT | 1020 |
| 10 | GCACTTAAAA | ATTCTTCTAA | AATGTCACTC | AATCTTGAAA | CTTTCATTTG | TTCATCATCA | 1080 |
| | GTTAACGTTT | CGCCATTACT | CATTTTTCCC | CAACCAGTTA | ATACATCTTT | GTCAATAACT | 1140 |
| | GGAATATAAG | TTTCAFATAA | ATCAATATGT | TTCACTTCAT | CATCAGGATG | TTGTTGTTGA | 1200 |
| 15 | TATGTTTCGA | TAAATGCTTT | ACCAGCCGCC | ATAGAATTTG | ATACCAGTTC | ATTAAAAGGG | 1260 |
| | TGTGCTGTAA | TATATAATAC | TTTTGCCATT | TGAAAATTCT | CCTCTGkTTC | TGTTATTTTC | 1320 |
| 5 <i>0</i> | TTAAGTATAA | TTATTATACT | CGATATAAAA | TTTAATATCA | ATCAAAATAT | TCAAATTACC | 1380 |
| | ATCATTTTCT | TCATCTATAT | nTGGCAGTAC | TACTAAAGTA | TGAGTGCATT | TAATTATGAa | 1440 |
| | ATAGTTGATT | TAGAATATAT | ACTTAATACC | СВВВВТВТВТ | GAAGGATGGA | TOCONCENTO | 1500 |

| | ATTATTTATA | TAGATGACAT | TCAAAAATGG | TTTAACCAAT | ATACCGATAA | ATTGACACAA | 1620 |
|------------|--------------|-------------|-------------|------------|------------|------------|------|
| | AATCATAAAG | GACAAGGACA | CTCAAAATGG | GAAGACTTTT | TTAGAGGGAG | TCGGATTACT | 1680 |
| 5 | GAGACTTTTG | GTAAATATCA | ACATTCACCA | TTTGATGGTA | AGCATTATGG | CATTGATTTT | 1740 |
| | GCATTGCCAA | AAGGTACACC | AATTAAAGCG | CCGACGAATG | GTAAAGTAAC | ACGTATCTTT | 1800 |
| 10 | AATAATGAAT | TGGGCGGCAA | GGTATTACAG | ATTGCCGAAG | ACAATGGAGA | ATATCACCAG | 1860 |
| | | | | | ATCGAGTCAA | | 1920 |
| | | | | | GCGCACATTT | | 1980 |
| 15 | AGAATGAAGG | GTGGCGTAGG | TAATGCATAT | GCAGAAGATC | CAAAACCGTT | TATCGATCAG | 2040 |
| | TTACCTGATG | GGGAACGTAG | CCTATATGAT | TTGTAGTTAT | AGAAGGGTGC | CCGCAGTCTA | 2100 |
| | AAAAATTAAG | CAATCATTGT | GTGAGTATGA | TACTTACATA | ATGGTTGCTT | TTTTCAATGA | 2160 |
| 20 | AAATCGTAAT | GCTAAGTCAT | ACTTGTTTGA | TTTAGATATT | ACTTAAAATG | TAAGACAAGG | 2220 |
| | TTGTTAGCAT | TGGCAGTGAA | ATATCGCACA | TAAAAAACAT | TATTGTCACA | CTAGAAAATA | 2280 |
| 2 5 | GTTGTGCACT . | ATATCAATTT | TCTGTATAAA | AGTTTAATTC | TGACAGTAAT | GTAAACGTTT | 2340 |
| .5 | ACAATTTATG | ATTGACATTA | ATAATGACTG | AATATATGAT | TTATGTAAGT | ATTTGTGCAA | 2400 |
| | CGTTTTCACA | AAGTGTATTG | САСаАуСААА | CTGtAAACaA | aGTATGGGGg | GCCATAACAT | 2460 |
| 30 | GGCAGAACTA I | AGTTAGAGCn | TATTAAAA | | | | 2488 |
| | (2) INFORMA | TION FOR SE | Q ID NO: 10 | 8: . | | • | |

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4093 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

40

45

50

55

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 108:

TTTTCTTTAT TTCAAMCTGT ATATTAATGA TGTCACTTCA TTTGATACGA TTCTTGATAA 60

CCTATTCAAA ATTCCGCCAA ATAACATAAA TATTATATAA ATGCCGATAC TTTTAATCAT 120

TTTCTACTTT TTCTTCGATA CGGAAACTTG TTTTCGAATT GAACACTTCA CCAGCTTTTA 180

AAATTGACGG TGCTTTTTCA CCATATAAAT TAATATCATT TGGTAAAAAAT TGTGTTTCTA 240

ATGTAAAGCC AGAATGTGGT TTATAAATAT TAAATGGACT ATCCCACTCA TCAGGCTGGT 300

TAAAAGTAAA GAACACAACA TGAGGCATAT CTGTATCGAC CTCTAACATA AATTCATGAT 360

TTTCAACATA CATTTTATGT TCACCAACTG TAAATGGGTG ATCGAGACCA CCAAAACGTG 420

| | TATCTTCAAA | CACTTCATGT | AAATCTAGAA | TATCACCTGT | AACAATATTT | CGCTCATCTA | 540 |
|----|------------|------------|------------|------------|------------|------------|------|
| _ | ATACATACAT | ATCTAATTGA | TTACTTGAAA | TGCGATGATT | ATCAACGACA | TTATTATCTC | 600 |
| 5 | GATTCAAATT | GAAGTACACA | TGATTCGTAG | GACTAAACAA | TGTGTCTTCT | GATGCAACTG | 660 |
| | CTTCGTATTC | AATCGACCAT | TGGTGATCCG | CATCATAAAT | ATGTGTAATC | GTCACATCGA | 720 |
| 10 | TATCACCCGG | GAAATGATCA | TCAGCTGATT | TCAACACCGT | СТТАААТАТА | ACTTTAATTT | 780 |
| | GAGCAATTTC | ATTTCTAATT | TCATAATCAA | ATAACTTATT | GTCCAAACCA | TGACATCCAC | 840 |
| | CATGTAAATG | ATGTTCACCG | TIGITITITT | CTAACTGATA | TTCTTTACCT | TTCAACTTAA | 900 |
| 15 | ATTTAGCATT | ATCAATTCTA | CCGCTATATC | TTCCTATAGA | AGCACCAAAT | TTAAAAGGAT | 960 |
| | TACTATGATa | AAATTCATCC | GCTTCAACAA | CATTTCCAAG | AACAATATTA | TTATCATGAT | 1020 |
| | ATTTCCAAGA | CACTACTCTT | GCTCCATAAT | TCGTAAAAAT | AATTTTAGTT | TCATCATTAT | 1080 |
| 20 | CAATTTTGAT | TAAATCTACA | CCTTGTCTTT | GGTGCTCAAC | TTCAACTATC | ATTTTTACTT | 1140 |
| | CTCCCTTCTA | ACCACAAGTG | TTCAAGCTCT | GCTGGGTAGC | AACATTACTA | AAACACCTAC | 1200 |
| 25 | AATACAAATG | ATTGCACCGA | TAACATCATA | TTTATCTGGC | ATTTGTTTAT | CTACGACCAT | 1260 |
| | CGCAAAAATC | AAACTCATGA | TGATAAATAC | GCCACCATAT | GCTGCATATA | CTCTTCCGAA | 1320 |
| | TGATGGAAAT | GATTGAAATG | TCGCAATGAC | ACCATATAAC | ATGAGTATCG | CACCGCCTAT | 1380 |
| 30 | TAGCCCAACA | AGTGAAGACT | GTCCTTCCCT | AAGCCACAGC | CAAATCAGGT | ATCCCCCACC | 1440 |
| • | TATTTCACAT | AAGCCAGCTA | ATATAAATAT | AAAAATCGGA | TATAACATGA | AATCACTCCA | 1500 |
| | TCACACATTT | GCTATCAATA | ATCTATCGGC | TACATATCAT | TTGTTTACAT | TTCTTCTTAC | 1560 |
| 35 | TTCACATTCC | CATTTTAAAA | AGTTCGTTTT | CACATTCATA | TTGTACACTT | TTTTAGACAT | 1620 |
| | TATTCTATAG | СТАААТАТАА | AAAAATAAGA | GTAACACGCT | TTCATCATCA | TTTTATATGA | 1680 |
| | TAAATGTGTG | TCACTCTCAT | CAATTTTATT | TTTTAAATAC | ACGTTTCATT | GAATTAAATA | 1740 |
| 40 | AGCCACGTTC | aaatgtaagt | ACTGAATCTT | TATATGTTTT | AATTGCAATC | CATATCAAGA | 1800 |
| | CAGCTACCAT | TACAATTGAG | ATTAAAGAAC | TTAAGATGAC | CTCATATATT | TGAAGCCCTG | 1860 |
| 45 | AAGTTTGAGC | GCGTACAACT | AATTGAAATG | GCGCTAAAAA | CGGAATATAA | CTTGTGATTA | 1920 |
| | AAGCAAGTTG | TCCATCAGGA | TTATTTATCG | TGAATATCGC | GATATAAAAT | GCAATCATAC | 1980 |
| | CAAGTAATGT | CAGTGGCATC | AAAGATTGAT | TTAAATCTTC | TATTCTAGAT | GTTAATGATC | 2040 |
| 50 | CGAGGATGGC | TGCAAGTAAT | ACATACGCCG | TAATTCCAAC | AATACTACTT | ATAATTCCGA | 2100 |
| | CAATAATAAT | TTGCCAAGAC | AATTGATTCA | TTTCCACGTT | AAAACCTTGT | AGCAAGTCTT | 2160 |
| | TTAAGTCAAA | GGCAAAAATG | CATATAACTG | CCATCAATAC | AATTAAAATA | ATCTGAGTCA | 2220 |
| 55 | | | | | | | |

| | TAATAATCAT | TTCAATGACA | CGCGATGTTT | TCTCACTAGO | AATTTCCATA | GCTATTTGAG | 2340 |
|----|------------|--------------|------------|------------|--------------|------------|------|
| | ATGCATAATI | TAAAACAATG | AAGAACATTA | GAAAGATAAT | GCCATmaGcT | AAAGCATAGT | 2400 |
| 5 | TGAAAATCTT | TTGTCCTTCT | GATACTTTAT | CGACTTCATO | ATTAGAAATC | ACCTTATTAT | 2460 |
| | CAACTTTACT | TTGTGCTTGT | AATTTTTGTA | AGTCTTCTTT | GTTGATATTT | AATTCCCCGG | 2520 |
| 10 | CTACCATATT | TGTTTGAATA | GCTGTAAGCA | GTGCTTGTAC | TTTTTGTGAA | TCTTCATGAC | 2580 |
| | TTACTCGCTT | CTCACTAATG | ATTGTCCCTT | GTAACGTGCG | ATTTTGATTC | ACCTTGATAA | 2640 |
| | TATAAGCTTT | ATCAAGTTTA | TGTTTTTTA | CTTCTTTTTC | AGCATCTTCT | ATAGAAACTT | 2700 |
| 15 | TAGTAAACTT | AGCATCACTA | TGAAATGTAT | TCGCCTGTTG | CTTGAAAACC | TTATAGATTT | 2760 |
| | GTTCATTCGG | TGCTGCTACA | CCAATTTTAT | CTGGACCATC | ATCAAACATG | TTAATAATCT | 2820 |
| | TATCAATGTT | AGATAGGCCA | ATCATTAAGG | CAGCAATAAT | AATCATAAAA | ATTACAAATG | 2880 |
| 20 | ATTTAGCTTT | AATTTTTTTG | ATATATGTCA | AAGTAAATGT | CGCCCAAAAC | TTATGCATCC | 2940 |
| | TTGCCACCAA | CCTTCTCAAT | GAATATATCT | TGTAATGATG | GTTCTACAAC | TTGGAATCGT | 3000 |
| 25 | TTAACATAAC | CTTGATGTGC | CACAACTTGA | TAAATATCTT | TGGCTACGTC | TTCATTCTCA | 3060 |
| 25 | ATCGTCAACT | GAAGACCTTG | CTTCATGTTT | TCACTATGAA | TGATGCCTCT | AATGTTTGTT | 3120 |
| | AAATCTGGTA | GTGTTGTTTC | TGATTCAAÍG | ACAACTTTCT | TGTTACCATT | AGATGCACGT | 3180 |
| 30 | ACATGATTGA | TATCACCAGA | AACAACAAGT | TGACCTTTAT | CTAAAATACA | AACATCATCA | 3240 |
| | CATAATTCTT | CAACATGCTC | CATACGGTGA | GAACTATAAA | CGATTGTACT | GCCCCAATCA | 3300 |
| | TTTAAGTCTT | TAACTGCTTC | TTTTAATAAC | TCAACATTAA | CTGGGTCTAG | ACCACTGAAA | 3360 |
| 35 | GGCTCATCTA | ATATTAGTAA | TTCTGGTTTA | TGTAACATAC | TTGCTAACAG | CTGAATTTTT | 3420 |
| | TGTTGATTCC | CTTTTGATAG | ACTATCAATT | CGTTTTTTGC | GGTTTTCAGT | AATATCAAAA | 3480 |
| | CGCTCAAGCC | AATACGATAT | TTGCTGTTGT | ATTTCTGTTT | TTGACATTCC | CTTTAAAGTT | 3540 |
| 10 | GCCAAATATT | TCAATTETTC | TTCAACTGTC | AATTTCCCAT | GTAAACCGCG | TTCTTCCGGT | 3600 |
| | AAATAACCAA | TACGATTGTA | CATTGTTTTA | TCTAGTTTTT | TACCGTTATA | CGTrrTGTGT | 3660 |
| 15 | CCTTCAGTTG | GTTCACTTAA | GCCTAAAATC | ATACGAAATG | TCGTTGTTTT | ACmTGCACCA | 3720 |
| | TTTCTTCCTA | GAAAACCTAA | CATTTTACCT | GATTCTAACT | TTAATGAAAT | ATCATTTACT | 3780 |
| | GCCGTCATCT | TGCCAAAACG | TTTCGTAACA | TGTTCAATTA | CAAGTCCCAT | ACTTTGCCTC | 3840 |
| 50 | CTAAAAAnAT | ATGTATTTAT | СТТААТАТАА | CATTTCCATT | CTCTATAAAT | GCAATATTTT | 3900 |
| | TAAAATGAAT | TTATTTTTAA . | AATTTCTGAA | ATTGAAAAAT | TTAAATAGTG | CCATTTTTGC | 3960 |
| | ATGTTAAGTA | TCATTAGCAC | TAGATATGTT | TTTTCCATGC | CTTTATTGCC : | TTATTTGTAA | 4020 |

626

| | CTTnCCGGTG TTT | 4093 |
|-----|---|------|
| | (2) INFORMATION FOR SEQ ID NO: 109: | |
| 5 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 17846 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 10 | (b) Topologi: Timear | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 109: | |
| 15 | TGCCAAACTA CCTTTTGACA GTCGTTGCTG TACTTCAGGA TGATCAATCA CATATNTTAC | 60 |
| | TTTATCAAAT AGGGCATCTT CATCATTTTT AGTAATTAAA TAACCATTGA AATCTGAAGT | 120 |
| | AATCAGTTCG TTAGGTCCAT ATTTAATATC ATAACTAATA ACTGGAACAC CATGTGCTAA | 180 |
| 20 | AGATTCAAGT AGCGCTAAAG AGAAACCTTC CATGTTACTT GTTATTAAAC TCAAATAGGC | 240 |
| | ATCGCTATAT TCTTGGTCTA GATTGCTTAA AAAGCCGCGT AAGTAAACAT GATTTTCCAA | 300 |
| | TCCATATTTT TGTATCAATT CATTTAATTT TTTACTTTCA GAACCAAAAC CATACATATG | 360 |
| ?5 | AAGCTCTATT TTTGGGACAT ACGATACTAA GCGTTTAATT AATTCAATTT GTTGATGTAA | 420 |
| | TIGTITITCA GGTGAATAAC GAGCAACGGA AATTAATTTA ACACTGCGCT GATCTAATGT | 480 |
| 80 | TTGGACTGGT GTATCAATTG TTTCACTATA GCCGACAGGA ATATTAACAA CTGGAATAGT | 540 |
| | ATGGTTAATA CGTTTTTCAA CATCTAATTT TTGCTGCTCA GTAGAAACGA TAATTGCACG | 600 |
| | ATATCGAGAT AAATTTTCAA ACATCGCTTT ATATACATTT TTAAATGGCG ATGAATCTAA | 660 |
| 35 | TGCATCAATA TTTTTAATGT GTGTACTGTG AAGCACAGCT ACTACTGGGA TTGACTCAGG | 720 |
| | CGTTAAGTTG AAAATAGGTG CTGTGTACAC ATTACGATCA CTGAAAAATA AATCCCCATG | 780 |
| | TTGATATAGT TGTTTAATGA AAAATGCGCC TAATTCCGTT TCATTATTAA AGAAATATTG | 840 |
| 10 | TTTGTTAGCA TAGTAAACAA TAATTTTTTG TACTTCTGGT TTGCCATCCT TGTAAGAAAA | 900 |
| | ATACTTTTCT AATTTTGTGT CACCTTCTGG ATTATAGAAA AATTCACATA ATGTTTGTTG | 960 |
| 15 | TTTATCAACA AGAATCCTAC TACAACTTAA AAAGCCACGC ACATCATAAA AATCACGTTT | 1020 |
| .5 | TACTTETCGT CTTTGACTAT CAAAATGATT TACATAATCT AATATACGAT ATTTAGGATC | 1080 |
| | TTGAAAATGG GCATACATTA AGAAACGCTC TTGATCATAT ATTCTAAAGT CATGACTATT | 1140 |
| 50 | TTCAACATGT TTTAAAGTAT AATGACATTC ATCAGTCCAA TACGACAACC AGTCAAATGG | 1200 |
| | TTCATTGCGT TCTAAATATG TTGCTTCTTG GAAGAAATCA TACATATTAA TATAGTCAGA | 1260 |
| | ACTAGTAATA TAATTTTGGG CATTTCTATA TAAATATCTA TTCCATGACA GAAATACACA | 1320 |
| : 6 | | |

| | CCCAGTTAAA | TTAACACCTA | AACTATTACC | TACAAAATAA | TTCATTTACA | ACACCACTTA | 1440 |
|----|--------------|-------------|-------------|------------|-----------------|------------|---------|
| | TATCTATTTT | TTATAATTAT | ATCACATAAT | ATTTAATTAC | TTCTTTTAAC | TGGAAGATGT | 1500 |
| 5 | GITTATTTAT | AAAACAACAA | ATTTTGATAT | TTATAATGAT | AGTAGTTATT | CAATCACTAC | 1560 |
| | GACCCAATAT | ATCATKGTAG | AGCTTAGGAT | ATTGATTTAT | GACTCAGGCA | CATCAAATGa | 1620 |
| 10 | GAGGATTTAT | AAArGAGATA | TACAACTCTA | GAAGGTATAA | TAAAAACGCG | CAACTAATGT | 1680 |
| | TACGCGTTTG | AATTAATCAT | ATGATATTAT | TTGCGATACT | TTAATTTAGC | GAAAgcATCA | 1740 |
| | TGTTGATGGA | TAGACTCTTC | ATTACGACAT | TCGATATCGA | AACCGTCTAA | CCAATCAAAT | 1800 |
| 15 | TCAACTAAGT | CCGCGGCAAT | TAAACGAATT | AAGTCTTCGA | CAAAACGTGG | ATTTTCATAT | 1860 |
| | GCACGCTCTG | TCACACGTTT | TTCATCAGGA | CGTTTTAAAA | TAGGGTATAG | AATTGAACTT | 1920 |
| | GCATTAGCTT | CCATTGCATC | TAAAATTTTA | TTTTTATAGT | CATCAACTAT | GTCTTGATCT | 1980 |
| 20 | TTATTAATAT | ATGTTTTAAC | AGTGACAACA | CCACGTTGGT | TGTGCGCTGA | ATACTCACTT | 2040 |
| | ATTTCTTTTG | AACAAGGGCA | TAGCGTTGTG | ACAGTTGCTT | CAATAGTAAG | TTCTTTACGT | 2100 |
| | GTAnCTTTAT | CACCGTCAAT | TGCTAATCCA | TAAGTGACAT | CGGCATTACC | AACTGCTTTA | 2160 |
| 25 | ATATTTGTGG | TTGGACTATA | GCGATCAAAG | AACCATTTCC | CAGAAACATC | AACGCCTGCC | 2220 |
| | GCATTTTGTT | TCATATTCGT | TTGTAAAGTG | CGTAACACCT | GATAAAGTGT | ATTAAATTCA | 2280 |
| 30 | AGTTCAATAC | CATTATCATA | GTGCTTTTCA | ACACTTTCGA | TTATACGGCT | CATATTAATA | 2340 |
| | CCTTTTTCGT | CTTTTGTTAA | ACTTGTTGAA | AAACTAAATG | TGCCAGCTGT | TTGATACTGG | 2400 |
| | TCAACAAGTA | CAGGGTACAC | TAAGTTTTTA | ATACCAACTT | CTTCTATTTC | AAATAAAAA | 2460 |
| 35 | TCTTTATGTG | TACTTTGTAA | ATCTGTCATT | TCGTTCTTAG | TAGTAGGTTT | CGTGCCTTCA | 2520 |
| | ATAGGATCTA | CGGAACCAAA | GTGTTTCCAA | CGACCTTCTC | GTGTCGATAA | ATCAAATTCA | 2580 |
| | GTCATTTTTT | TCCTCCGTTA | AGATTTAAAG | TGATATGTCC | AATATGGTTC | GACTGTTAAA | 2640 |
| 40 | AAGCTGTGTT | GTTTACCATC_ | GATTTCAGGA | CTTGCTAATT | GTTTTAAAAA. | TGGACCTGTT | -2700 - |
| | TGAGAAGCAT | GTGCTTCAAA | TGCCTTAATT | TTAAGTTCTT | TAAAATCTGT | AATATCATTT | 2760 |
| 45 | TGAATATCAG | GTTCTCCAAG | AGCTTCGGTT | GCATCATTAC | TGAACGCAAC | TAAAGTTAAA | 2820 |
| | CGAGGGCGTT | CTTCTTTAGG | CATGCGTTCA | ACCGTTCGAA | TTACAGCGTC | TGCTGTTGCT | 2880 |
| | TCGTGATCAG | GATGTACTGC | ATATCCAGGA | TAAAATGAAA | TAATCAATGA | TGGATTTGTA | 2940 |
| 50 | TCATCGATTA | AAGATTTAAT | CATACCATCT | ATATGTTCAT | AGGGTTCAAA | TTCGACAGTT | 3000 |
| | TTGTCACGTA | AACCCATTTT | TCTTAAATCA | GTAATACCGA | TAACTTTACA | AGCTTCTTCT | 3060 |
| | AGTTCACGCT (| CACGAATACT | теста атеат | TCCCCTCTTC | CN N N TO CO CO | | |

| | TAATTTGCTA | ATGTGCCTGC | AGATGAGAAG | GTTTCATCAT | CAGGATGTGG | AAATATTACT | 3240 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | AATACATGTC | TTTCGTCAGT | CATGTTGATG | CCTCCTCTAT | AAATTAAATG | GTCGCTCACT | 3300 |
| 5 | AATTTGAAGT | GCTGCAGCGA | GTTGACCTTC | GTAATTAAAA | CCTGCAATTA | AAAATTCATC | 3360 |
| | ATGCTCATTG | ACCTCAAAAT | GCGTTAGACC | TTGTACATAA | ACCCAACCAC | CATTTGATAG | 3420 |
| 10 | TTTAAGACCA | ATGCGATAAG | GTTCTTTATT | ACCACCTTTT | AGTTGTGCAT | GCGTATATGT | 3480 |
| | TATTTGTATG | TTTCTTAAAA | AAGTACCAGC | ATTAAAAACA | CGTTGATCGA | AATGGTTCGC | 3540 |
| | ATAGGCCCCA | TTTGTCGTTT | CAACATGCAG | ATACACAGGT | TTATGTTCAA | AAGAAGCAAG | 3600 |
| 15 | TAAATCTATA | ACTTCTTGTT | CTTTAATTGG | TTCCAACACG | TTCACTCCTT | ACACTATCAA | 3660 |
| | TGTGTTTATC | TTTCTATTTT | ACTAAAAACT | ATTCGATAAT | TGTATACGAT | TGCTCAATTA | 3720 |
| | TTTATAAATT | AATTTTCATG | AAGGGTAATT | ACTCAGGATT | ACGTAATCAT | ACAGCATTAG | 3780 |
| 20 | TTTTTTACTT | TTAAAAATCA | AAAATTTGTT | GGAATTTGAA | AAGTGTTAAA | CATTAAAAAT | 3840 |
| | GATGCTATAT | TAATGGTGTA | TGAATGAATT | CATAAGTTTT | TAAAATGTAT | TAAATTTGTG | 3900 |
| 25 | GAGGCATGTA | AACAATGAAA | GTATTAAACT | TAGGATCGAA | AAAACAAGÇA | TCATTCTATG | 3960 |
| 2 5 | TTGCATGTGA | GTTATATAAA | GAGATGGCAT | TTAATCAGCA | CTGTAAACTA | GGTTTAGCAA | 4020 |
| | CTGGTGGTAC | AATGACAGAT | TTGTATGAAC | AACTTGTTAA | GTTGTTAAAT | AAAAATCAGT | 4080 |
| 30 | TAAACGTAGA | CAATGTATCC | ACGTTTAATT | TAGACGAATA | TGTAGGTTTA | ACCGCATCAC | 4140 |
| | ATCCGCAAAG | TTATCACTAT | TATATGGATG | ACATGCTTTT | CAAACAATAT | CCTTATTTTA | 4200 |
| | ATAGAAAGAA | CATTCATATT | CCAAATGGAG | ATGCCGATGA | TATGAATGCG | GAAGCGTgCA | 4260 |
| 35 | AAATATAATG | ACGTTTTAGA | ACAACAAGGT | CAACGTGATA | TTCAAATTTT | AGGTATTGGT | 4320 |
| | GAAAATGGTC | ATATTGGATT | TAATGAACCT | GGTACGCCGT | TTGATAGCGT | TACTCATATC | 4380 |
| | GTTGATTTGA | CTGAAaGTAC | TATTAAGGCT | AATAGTCGAT | ATTTTAAAAA | CGAaGATGAT | 4440 |
| 40 | GTTCCAAAGC | AAGCCATTTC | GATGGGACTT | GCTAATATTC | TTCAAGCCAA | ACGTATCATT | 4500 |
| | TTACTCGCAT | TTGGTGAAAA | GAAACGTGCT | GCTATTACAC | ATTTATTAAA | TCAGGAAATT | 4560 |
| 45 | TCTGTTGATG | TTCCAGCCAC | ATTACTTCAC | AAACACCCGA | ATGTTGAGAT | ATATTTAGAC | 4620 |
| +0 | GACGAAGCTT | GCCCGAAAAA | TGTTGCGAAA | ATTCATGTCG | ATGAAATGGA | TTGATTGCAA | 4680 |
| | TGTTTAATTA | AGAAATGCCT | CGGGAAAGGT | TCCAATAGAA | AGATAAAAAG | CATTGGAAGG | 4740 |
| 50 | ATGATTTTTA | GTGGAATTAC | AATTAGCAAT | TGATTTATTA | AACAAAGAAG | ACGCGGCTGA | 4800 |
| | GTTAGCAAAT | AAAGTAAAAG | ATTATGTAGA | TATCGTAGAA | ATCGGTACGC | CAATCATTTA | 4860 |
| | CAACGAAGGT | TTACCAGCAG | TTAAACATAT | GGCAGACAAC | ATTAGTAATG | TAAAAGTATT | 4920 |

| | CGCGGATGTA | ATTACAATAC | TAGGTGTTGC | AGAAGATGCA | TCAATTAAAG | CAGCTATTGA | 5040 |
|-----------|--------------|------------|------------|------------|------------|------------|------|
| | AGAAGCTCAT | AAAAATAATA | AACAATTACI | AGTTGATATG | ATTGCTGTTC | AAGATTTAGA | 5100 |
| 5 | AAAACGTGCA | AAAGAACTAG | ATGAAATGGG | TGCTGATTAT | ATTGCAGTAC | ACACTGGTTA | 5160 |
| | TGATTTACAA | GCAGAAGGGC | AATCACCATT | AGAAAGTTTA | AGAACCGTTA | AATCTGTTAT | 5220 |
| 10 | TAAAAATTCT | AAAGTTGCAG | TAGCAGGTGG | AATTAAACCA | GATACAATTA | AAGATATTGT | 5280 |
| | CGCTGAAAGT | CCTGATCTTG | TTATTGTTGG | TGGCGGAATC | GCAAATGCAG | ATGATCCAGT | 5340 |
| | AGAAGCTGCG | AAACAATGTC | GCGCTGCAAT | CGAAGGTAAG | TAATATGGCT | AAATTTAGTG | 5400 |
| 15 | ACTATCAATT | AATTCTAGAT | GAATTAAAGA | TGACTTTGTC | ACATGTTGAA | GCGGATGAGT | 5460 |
| | TTTCAACTTT | TGCATCCAAA | ATACTACATG | CTGAACATAT | ATTTGTAGCT | GGCAAAGGAC | 5520 |
| | GTTCAGGATT | CGTGGCGAAT | AGTTTTGCAA | TGCGCTTAAA | TCAGCTCGGC | AAACAGGCAC | 5580 |
| 20 | ATGTTGTTGG | AGAATCAACG | ACACCTGCGA | TTAAGTCGAA | TGATGTATTT | GTAATTATCT | 5640 |
| | CTGGTTCAGG | TTCCACGGAA | CATTTAAGAT | TATTAGCAGA | CAAAGCAAAA | TCAGTAGGTG | 5700 |
| | CTGACATCGT | ATTAATTACT | ACAAATAAAG | ATTCTGCAAT | AGGCAATCTA | GCTGGGACGA | 5760 |
| 25 | ACATCGTTTT | GCCTGCAGGT | ACAAAATATG | ATGAACAAGG | CTCGGCACAA | CCATTAGGAA | 5820 |
| | GTTTGTTTGA | ACAAGCATCT | CAATTATTTT | TAGATAGTGT | TGTAATGGGA | TTGATGACTG | 5880 |
| 30 | AAATGAATGT | TACGGAACAA | ACGATGCAAC | AAAATCATGC | TAATTTAGAA | TAAAATAAAG | 5940 |
| | ATAGTCGATA | ATATGATGCC | TAGGCAGAAA | TATTATCGAT | TATTTTTTTA | TTAAATAAT | 6000 |
| | AAATTATAGT | ATAATATCAA | TAATAAACGA | ATAGGGGTGT | TAATATTGAA | GTTTGACAAT | 6060 |
| 35 | TATATTTTTG | ATTTTGATGG | TACGTTGGCA | GACACGAAAA | AATGTGGTGA | AGTAGCAACA | 6120 |
| | CAAAGTGCAT | TTAAAGCATG | TGGCTTAACG | GAACCATCAT | CTAAAGAAAT | AACGCATTAT | 6180 |
| | ATGGÉAATAC | CTATTGAAGA | ATCATTTTTA | AAATTAGCAG | ACCGACCATT | AGATGAAGCA | 6240 |
| 40 | GCATTAGCAA | AGTTAATCGA | TACATTTAGA | CATACATATC | AATCTATTGA | AAAGGACTAT | 6300 |
| | ATTTATGAAT | TTGCGGGTAT | AACTGAAGCC | ATTACAAGTT | TGTATAACCA | AGGGAAAAA | 6360 |
| 45 | CTTTTCGTGG | TGTCTAGTAA | GAAGAGTGAT | GTATTAGAAA | GAAATTTATC | GGCTATTGGA | 6420 |
| | TTAAATCACT | TGATTACCGA | AGCTGTTGGA | TCCGATCAAG | TAAGTGCATA | TAAACCAAAT | 6480 |
| | CCTGAAGGCA | TACACACAAT | TGTGCAACGC | TACAATTTAA | ATAGCCAACA | AACGGTGTAT | 6540 |
| 50 | ATTGGTGATT | CAACGTTTGA | TGTTGAGATG | GCACAACGTG | CTGGTATGCA | ATCTGCAGCT | 6600 |
| | GTCACTTGGG (| GTGCACATGA | TGCAAGGTCA | TTACTTCATT | CAAATCCGGA | TTTTATTATT | 6660 |
| | AATGATCCAT | CAGAAATTAA | TACCGTATTA | TAAAACTTGT | TAAAACAGAG | AATACCATGG | 6720 |

| | ATTTAAAATA | AATATTTATT | ` AAACATTATG | AATTTTTAAA | GAGTAATGTC | TGACTCGTTG | 6840 |
|----|------------|------------|--------------|------------|------------|------------|------|
| _ | ATAATTTATT | TTIGTAAAAA | TAAATTAAAG | TAATGACAAA | GTTATTGAAG | TAAATTGAGT | 6900 |
| 5 | ATAAACATTT | AAATACGATG | TCGAAAATGG | CGATAGCATA | TCACTTACAT | GAAGTTGTGT | 6960 |
| | GCLATCGCTA | TTTTTAGTTA | TAATTCCAAA | AAGTTAATCG | TTCGATGATT | TAAGAATTAT | 7020 |
| 10 | TATTGTTTAA | TTCAAATGTA | TGAGGGTATA | AAATCATTGA | ATTTAATTCG | ATAAAGCGAA | 7080 |
| | ATTTTTGAAC | AAACATACTT | TTGTATTTAT | ATAAAAGTTT | AAATTCTTAT | AAATTTGACA | 7140 |
| | AAACTAATTA | ACTCCGTATA | ATTATGAAAC | ATACAAGAGG | GAGTGTATGA | ATTCATGGAT | 7200 |
| 15 | TTTAATAAAG | AGAATATTAA | CATGGTGGAT | GCAAAGAAAG | СТАААААААС | CGTTGTTGCA | 7260 |
| | ACCGGTATCG | GTAATGCAAT | GGAATGGTTC | GATTTTGGTG | TCTATGCATA | TACAACTGCG | 7320 |
| | TACATTGGAG | CGAACTTCTT | CTCTCCAGTA | GAGAATGCAG | ACATTCGACA | AATGTTGACT | 7380 |
| ?0 | TTCGCAGCAT | TAGCCATTGC | GTTTTTATTA | AGACCAATTG | GTGGTGTCGT | ATTTGGTATT | 7440 |
| | ATTGGTGACA | AATATGGACG | TAAAGTTGTA | TTAACATCTA | CAATTATTTT | AATGGCATTT | 7500 |
| | TCAACATTAA | CCATTGGATT | ATTGCCAAGC | TATGATCAAA | TTGGACTTTG | GGCACCAATA | 7560 |
| ?5 | CTATTATTGC | TTGCAAGAGT | ACTACAAGGG | TTTTCAACAG | GTGGAGAGTA | TGCGGGGGCA | 7620 |
| | ATGACATATG | TTGCCGAATC | ATCTCCAGAT | AAGCGTCGTA | ACTCATTAGG | TAGTGGACTA | 7680 |
| 80 | GAAATTGGGA | CATTATCAGG | TTACATAGCT | GCTTCAATTA | TGATTGCTGT | ATTAACATTC | 7740 |
| | TTTTTAACAG | ATGAACAAAT | GGCATCATTT | GGTTGGAGAA | TCCCATTCTT | ACTCGGTTTA | 7800 |
| | TTCCTAGGAT | TATTCGGCTT | ATATTTACGT | CGTAAGCTGG | AAGAATCACC | AGTTTTCGAA | 7860 |
| 15 | AATGATGTTG | CAACACAACC | AGAAAGAGAT | AACATTAACT | TTTTACAAAT | CATCAGATTT | 7920 |
| | TATTACAAAG | ATATATTTGT | ATGTTTTGTA | GCTGTTGTAT | TCTTCaATGT | TACAAACTAT | 7980 |
| | ATGGTAACTG | CATATTTACC | AACCTATTTA | GAACAAGTTA | TTAAATTAGA | TGCAACGACA | 8040 |
| 10 | ACAAGTGTAT | TAATTACTTG | TGTCATGGCA | ATAATGATTC | CATTAGCATT | AATGTTTGGT | 8100 |
| | AAGTTAGCGG | Ataaaatagg | TGAAAAGAAA | GTATTTCTAA | TTGGTACTGG | TGGGCTAACA | 8160 |
| | TTATTCAGTA | TCATCGCATT | TATGTTATTA | CATTCACAAT | CATTTGTTGT | AATAGTAATC | 8220 |
| 15 | GGTATATTTA | TATTAGGATT | TTTCTTATCA | ACTTACGAAG | CGACAATGCC | AGGGTCGTTA | 8280 |
| | CCAACGATGT | TTTACAGTCA | TATAAGATAT | CGAACTTTAT | CAGTAACATT | TAATATCTCT | 8340 |
| 50 | GTTTCGATAT | TTGGTGGTaC | GaCGCCATTA | GTkGCAmCaT | GGTTaGTTAC | GAAAACTGGA | 8400 |
| | GATCCATTAG | CmCCTGCGTA | TTATTTAACA | GCAATCAGTG | TTATTGGCTT | TTTAGTTATT | 8460 |
| | ACATTCTTAC | ATTTAAGTAC | AGCAGGAAAA | TCTCTAAAAG | GTTCGTATCC | AAATGTAGAT | 8520 |

| | GAACGTAAGA | ATTAGAGATT | TTAATAAAA | GTATAAATCA | ATCGTATATA | AGCACTTTAA | 8640 |
|-------------|------------|------------|------------|------------|------------|------------|-------|
| | | | TTTAAAGTGC | | | | 8700 |
| 5 | | | TAAACCCCCA | | | | 8760 |
| | | | AGCTTTGAAT | | | | 8820 |
| | | | GTATTTTGTC | | | | 8880 |
| 10 | • | | CAAACTTAAA | | | | 8940 |
| | | | TTCTCTATAT | | | • | 9000 |
| | | | TTTATCTAGG | | | | 9060 |
| 15 | | | TTCATTATAT | | | | 9120 |
| | - | | GAACGTAATT | | | | 9180 |
| | • | | ACATATCGAT | | | • | 9240 |
| 20 | | | TGAATCTTTG | | | | 9300 |
| | | | GCAAAGTATT | | | | |
| 25 | | | | · | | ATTGATTGAA | 9360 |
| | | | AATTGGAGAT | | | | |
| | | | ATTACAGCAC | | | | 9480 |
| 30 . | | | AAAAGCATAT | | | • | 9540 |
| | | | GATGTTAAAA | | | | 9600 |
| | | | | | | | 9660 |
| 35 | | | ATATGCGTTA | · | | | 9720 |
| | _ | | TCCTGAAAAG | | | | 9780 |
| • | • | | TGTTCCAACG | | | | 9840 |
| 40 | | · · | ATTTTTTAGC | | | | 9900 |
| | | | AAATGACATA | | | | 9960 |
| | | | | | | AGGTGTGCTA | |
| 45 | | | | | | AACTATTTGT | |
| | | • | | • | | TAATGATGAA | |
| 50 | | | | | | AACGGGACGT | |
| | | | | | | TGAACGCCTT | |
| | TTAACGCAAT | CTTCGAGCAT | TGATGAAGCA | ATTATCATCG | GTATTCCAAA | TGAGCGTTTT | 10320 |

| | CAATTTTTAA | AAAAGAAAGT | GAAaCgnTaT | GAAATTCCAT | CGATGATTCA | TCATGTAGAA | 1044 |
|-----------|------------|------------|------------|------------|------------|----------------|-------|
| | AAGATGTATT | ACACTGCAAG | tGGTaAAATT | GCTAGAGAAA | AAATGATGTC | GATGTATTTG | 1050 |
| 5 | AGAGGTGAAT | TATAATATGA | ATCAAGCAGT | CATAGTTGCA | GCTAAACGAA | CTGCATTTGG | 1056 |
| | GAAATATGGT | GGCACTTTAA | AACATTTAGA | GCCaGAACAA | TTGCTTAAAC | CTTTATTCCA | 1062 |
| 10 | ACATTTTAAA | GAGAAGTATC | CAGAGGTAAT | ATCTAAAATA | GATGATGTAG | TTTTAGGTAA | 10680 |
| 10 | TGTTGTTGGG | AATGGTGGCA | ATATTGCAAG | AAAAGCATTG | CTTGAAGCGG | GGCTTAAAGA | 10740 |
| • | TTCAATACCT | GGCGTCACAA | TCGATCGGCA | ATGTGGGTCT | GGACTTGAAA | GTGTTCAATA | 10800 |
| 15 | TGCATGTCGC | ATGATCCAAG | CCGGAGCTGG | CAAGGTATAT | ATTGCAGGTG | GTGTTGAAAG | 10860 |
| | TACAAGTCGA | GCACCTTGGA | AAATCAAACG | ACCGCATTCT | GTGTACGAAA | CAGCATTACC | 10920 |
| | TGAGTTTTAT | GAGCGTGCAT | CATTTGCACC | TGAAATGAGC | GACCCATCAA | TGATTCAAGG | 10980 |
| 20 | TGCTGAAAAT | GTGGCCAAGA | TGTATGATGT | TTCAAGAGAA | TTACAAGATG | AATTTGCTTA | 11040 |
| | TCGAAGTCAT | CAATTGACAG | CGGAAAATGT | AAAGAATGGA | AATATTTCTC | AGGAAATATT | 11100 |
| | ACCTATAACC | GTTAAAGGAG | AAATATTCAA | CACTGATGAA | AGTCTAAAAT | CACATATTCC | 11160 |
| ?5 | GAAAGATAAC | TTTGGCCGAT | TTAAGCCCGT | GATCAAAGGT | GGGACCGTTA | CCGCTGCGAA | 11220 |
| | TAGTTGTATG | AAAAATGATG | GTGCAGTTTT | ATTGCTTATT | ATGGAAAAAG | ATATGGCATA | 11280 |
| 30 | CGAATTAGGT | TTCGAGCATG | GTTTATTATT | TAAAGATGGT | GTTACGGTAG | GTGTTGATTC | 11340 |
| 50 | TAATTTTCCT | GGCATTGGTC | CAGTACCAGC | CATTTCCAAC | TTACTAAAAA | GAAATCAATT | 11400 |
| | AACGATAGAA | AATATTGAAG | TCATTGAAAT | TAACGAAGCG | TTCAGTGCAC | AGGTAGTTGC | 11460 |
| 35 | CTGCCAACAA | GCTTTAAATA | TTTCAAATAC | GCAATTAAAT | ATATGGGGTG | GTGCATTAGC | 11520 |
| | ATCAGGTCAT | CCATACGGTG | CAAGCGGTGC | CCAATTAGTG | ACTCGATTAT | TTTATATGTT | 11580 |
| | TGACAAAGAG | ACTATGATTG | CATCTATGGG | GATAGGGGGA | GGTCTAGGAA | ATGCAGCATT | 11640 |
| 10 | ATTTACTCGA | TTCTAACCAG | CGATTAAATG | TGTCATTTTC | TAAGGATAGT | GTGGCTGCAT | 11700 |
| | ATTATCAGTG | TTTTAACCAA | CCTTATAGAA | AAGAAGTACC | ACCATTAATG | TGTGCGTCAT | 11760 |
| | TATGGCCAAA | ATTTGATTTA | TTTAAAAAAT | ATGCAAATAG | CGAACTGATT | TTAACAAAAT | 11820 |
| 15 | CAGCAATTAA | TCAAACTCAA | AAGATAGAAG | TAGACACAAT | ATATGTAGGG | CATTTAGAAG | 11880 |
| | ATATTGAATG | CCGACAGACT | CGCAATATCA | CACGTTATAC | AATGGCTTTA | ACATTAACTA | 11940 |
| 50 | | ζ. | | | | AAGTAGAGAT | 12000 |
| 50 | GGAGTTTAAT | GAGATATGGA | TAAATGAATA | TTTGGCGCTC | GTAAATGATG | ATAATCCAAT | 12060 |
| | ACATAATGAG | ATTGTGCCAG | GACAATTAGT | GAGTCAAATG | ATCCTCATCC | CTD TCTC N TYP | 12120 |

| | ATTCATTGAA | CAACACGAAC | ACGAAATTAT | AGCAATTAAT | GACGATGGAG | AGATTAAAAT | 12240 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | AAAAATTTCT | TTGAGCACAA | AAAAATAACC | GATATTAGCT | GCATGAACGC | ATATTAATTA | 12300 |
| 5 | GGAGATGAAA | GGACAGCTAA | TATCAGTTAT | GTATTGTTAT | TATTATTGGG | AACAGAGATG | 12360 |
| | AATATAGGTT | ACGTTTCTTT | CTTTGCACGG | GGATGCATTA | АТСТААААТА | ATAATAACAA | 12420 |
| 10 | CTATATCAAT | GTTTAATAAA | TTCTGGATTA | TTGGAACGAT | TAGTCAATTT | AACTAACTTT | 12480 |
| 10 | CATATGATCT | ATATCGTCTT | GTAATAAAGA | GAGCAATTTG | AATATTTCAG | TATCACTAAA | 12540 |
| | TGAATCGTCA | CATTTAATTG | AAACATGCTG | AAACGTTTTG | GTTATAATTT | CATAAACTGG | 12600 |
| 15 | TGCGCCTTCA | TGGTGATACT | GTCGATAAAT | AATCATAACC | TATATTACCT | CCTTTGCTAC | 12660 |
| | TCTATGGTTA | TATTATAAAT | AACATTTTTA | TGTGTGACAT | CAACCTTAAG | TATCAACTTT | 12720 |
| | TTATCAGACA | TAGAACGTAT | GATTTACTAA | GACTATTTAT | GTATAAAAGT | ТСТАААТААА | 12780 |
| 20 | TATATATTTA | TAGAGTCGCC | TGGCAGTCAT | TTGGGaAATA | TAACATATAT | GATTAGAGAG | 12840 |
| ٠ | GCATCTATCG | CAAAAGAATG | ATAATGATAG | AGGTATTGAG | CATATAGATG | AGTTTAAGTT | 12900 |
| | CATCTTGAAA | ATAAAGGGTT | ATTTAGTCAT | AGATGTAGAT | GTATAGGAAA | TATTTGTATG | 12960 |
| 25 | TATTGTTCGA | TATGTATGAA | ATTTTCAATA | AAAGCTAATA | ACGCTTATAT | GTAACTTTCA | 13020 |
| | AATTTAAATT | ATATACAGAG | CATGATGATT | ATAAAAAAAT | AACCACATCA | CATAAATTGA | 13080 |
| 20 | GTTCATACCC | AATTTAAGTG | GTGTGGCTAA | TAATGTTGAT | TTATAGATGA | ACCGCCTAAT | 13140 |
| 30 | CGTTAAACCT | CTGTTACTTC | AACATCGATA | TGTTCAATAC | GGTTGTATGC | ACCGTGATCC | 13200 |
| | ACAGGACCAA | CAAAATCATT | CATTTTCCAA | CCGTTTTTAA | TAGCAGAAGC | GACGAAAGCT | 13260 |
| 35 | TTCGCGCTAA | TCACAGCTTC | TTTCGGTGAC | TTACCGTTAG | CTAAATATGC | AGTTGTTGCC | 13320 |
| | GCAGCAAATG | TACAACCAGC | ACCATGGTTA | TAACTTTGTT | GGAACATGTC | TGTTGTTAGT | 13380 |
| | TGATĀAAATG | TTTGACCATC | ATAGTATAAG | TCATACGATT | TATCTTGATC | TAAAGCTTTG | 13440 |
| 40 | CCACCTTTAA | TGATGACATG | CTGTGCGCCT | TŢATCAAAGA | TAATTGTTGC | AGCCTTTTTC | 13500 |
| | ATATCTTCAA | TTGAATTTAA | TTTACCTAAT | CCTGATAATT | GACCCGCTTC | AAATAAGTTT | 13560 |
| | GGTGTCACTA | CCGTTGCTTT | AGGTAGTAAA | TATTTAATCA | TCGCCTCAGT | ATTTCCAGGA | 13620 |
| 45 | TTAAGCACTT | CATCTTCGCC | TTTACAAACC | ATGACAGGAT | CTACTACAAA | ATATTGTGCA | 13680 |
| | TTAGATGCCT | CATATACTTC | TCCAGCACGT | TTGATTATCT | CCTCAGTACC | TAACATACCT | 13740 |
| 50 | GTTTTAATAG | CATCAGGTCC | GATTGATAAA | GCCGTTTCAA | GTTGTTTTTC | AAATACATCC | 13800 |
| 50 | ATTGGTAATG | GTGTAACATC | GTGTGACCAT | GTATCTTTAT | CCATAGTAAC | GATGGCAGTT | 13860 |
| | AAAGCGACCA | TGCCATACGT | ATCTAATTCT | TGGAACGTTT | TCAAATCTGC | TTGCATACCT | 13920 |

| | CACTCCTACA | TAATAATATT | GTATTCATCA | TATCATTTTT | AACCTAATTG | ATTATAAAAA | 14040 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | AGCATTCAAT | ATTTGATGAT | TGTTGAAATG | AATCATTCAT | ACTATTGTAA | CTTTTGAAAA | 14100 |
| 5 | TGTCATTCAC | TTTAGATAAG | TGTGATATGT | TAAAATATGT | CCTGAGGTGA | GATTGAATGG | 14160 |
| | AATGGTCGCA | AATTTTTCAT | GACATAACAA | CGAAACATGA | CTTTAAAGCT | ATGCATGATT | 14220 |
| 10 | TTTTAGAAAA | AGAATATTCG | ACTGCAATCG | TATACCCTGA | TAGGGAAAAT | ATATATCAAG | 14280 |
| 10 | CGTTTGATTT | AACACCGTTT | GAAAATATCA | AAGTTGTTAT | ATTAGGACAA | GACCCGTATC | 14340 |
| | ATGGTCCAÄA | CCAAGCACAT | GGATTAGCAT | TTTCAGTGCA | ACCTAACGCA | AAATTCCCTC | 14400 |
| 15 | CATCTTTACG | TAATATGTAT | AAAGAATTAG | CAGATGATAT | TGGATGCGTT | AGACAAACAC | 14460 |
| | CGCATTTACA | AGATTGGGCA | AGAGAAGGCG | TCTTGTTATT | GAATACAGTT | TTAACCGTAA | 14520 |
| | GACAGGGTGA | AGCAAATTCT | CATCGTGATA | TTGGTTGGGA | AACATTTACT | GATGAAATTA | 14580 |
| 20 | TTAAAGCAGT | GTCTGATTAT | AAAGAACATG | TTGTCTTTAT | TTTGTGGGGG | AAACCTGCAC | 14640 |
| | AGCAAAAAAT | AAAGCTTATC | GATACATCTA | AACATTGTAT | TATAAAATCA | GTGCATCCTA | 14700 |
| | GTCCACTGTC | TGCATATAGA | GGATTCTTTG | GATCAAAACC | GTATTCCAAA | GCGAATGCCT | 14760 |
| 25 | ATTTAGAGTC | AGTAGGAAAA | TCACCAATTA | ATTGGTGTGA | AAGTGAGGCG | TAGATGTTGA | 14820 |
| | ATAGAGAAAC | TTTAATAGCA | CGAATTGAGC | AAGAATTAGT | ACAAGCAGAG | CAGGCACAGC | 14880 |
| 30 | ATGACCATGA | CTTTGAAAAA | CATATGTATG | CCATACATAT | ATTAACATCT | TTATATGCTT | 14940 |
| 30 | CAACATCAAA | TACACCACAT | ATTGGTGAAC | AACAAATGAA | TCGTCGTATT | GCTAACCATA | 15000 |
| | ATCAAATGCC | ACAATCACAA | ATAACGCAGC | CAACTCATCA | AGTGACAGTT | GCTGAAATTG | 15060 |
| 35 | AAGCGATGGG | TGGTAAAGTA | AATACGCATT | CAGCACATCA | TCATAATAAG | TCATATTCAC | 15120 |
| | AACCTTCAAA | CCAACAACAA | AGATTAGCGA | CAGATGATGA | CATTGGCAAT | GGTGAATCCA | 15180 |
| | TATEIGATTT | TTAAAAAGCA | ACAATGAAAC | ATAATTACTT | AATAGCTTGT | TAAGTATGTA | 15240 |
| 40 | GGTTAATAAT | CAAGACGCAT | ATACTTTTAT | TCGAGTGTTC | GGATTTAAAC | ATTTATTAAT | 15300 |
| | ACTGAATTAT | ATAAGGAGAG | GTAGCAATGA | AATTATTTAT | TATTTTAGGT | GCATTAAACG | 15360 |
| | CGATGATGGC | TGTCGGTACA | GGTGCATTTG | GTGCGCATGG | TTTACAAGGA | AAAATAAGTG | 15420 |
| 45 | ATCACTATTT | ATCAGTATGG | GAAAAAGCAA | CGACGTATCA | AATGTACCAT | GGCTTAGCAT | 15480 |
| | TATTAATTAT | AGGTGTAATT | AGTGGTACAA | CTTCAATCAA | TGTTAACTGG | GCTGGCTGGT | 15540 |
| 50 | TAATATTTGC | TGGTATTATT | TTCTTTAGTG | GATCATTATA | TATTTTAGTA | TTAACTCAAA | 15600 |
| 30 | TTAAAGTTTT | AGGTGCGATT | ACGCCAATTG | GTGGCGTATT | GTTCATCATT | GGATGGATAA | 15660 |
| | TGTTAATCAT | TGCGACATTC | AAATTTGCTG | GTTAAATTTT | AAAACTTTAG | ATTACCTATG | 15720 |

| | TGGGTATAGA | ATACCTTCGA | GGTGAGTTTT | TATTTATGGA | AAAAAAGAAT | AAGCAAATAG | 15840 |
|------|------------|------------|------------|------------|------------|-------------|-------|
| | ATAGAGGCGA | TTTAAAACAA | AACCTATCTG | AAAAGTTTGT | ATGGGCGATT | GCATATGGTT | 15900 |
| 5 | CATGTATCGG | ATGGGGCGCA | TTCATCTTAC | CAGGAGACTG | GATTAAGCAG | TCAGGTCCGA | 15960 |
| | TTGCAGCATC | AATTGGTATA | GTTATTGGTG | CATTATTAAT | GATATTAATT | GCGGTTAGTT | 16020 |
| 10 | ATGGCGCATT | AGTAGAGAGA | TTTCCAGTAT | CAGGGGGCGC | GTTTGCCTTT | AGTTTCTTAA | 16080 |
| | GTTTCGGCAG | ATATGTGAGT | TTCTTCTCAT | CATGGTTTTT | AACTTTTGGT | TATGTCTGTG | 16140 |
| | TCGTTGCTTT | AAAtGCGACC | GCATTCAGTT | TACTAGTTAA | ATTCTTATTG | CCAGATGTCT | 16200 |
| 15 | TAAATAATGG | GAAACTATAC | ACCATTGCGG | GCTGGGACGT | TTATATTACG | GAAATCATTA | 16260 |
| | TTGCGACCGT | ATTACTACTT | GTATTCATGC | TAGTAACGAT | TCGTGGCGCA | AGTGTATCTG | 16320 |
| | GATCATTACA | ATATTATTTC | TGTGTGGCGA | TGGTAATCGT | CGTATTATTG | ATGTTCTTTG | 16380 |
| 20 | GTTCATTCTT | TGGTAATAAT | TTTGCACTTG | AAAATTTACA | ACCGTTAGCT | GAACCTAGCA | 16440 |
| | AAGGATGGTT | AGTGTCTATT | GTGGTTATTG | TATCCGTGGC | ACCATGGGCA | TATGTTGGAT | 16500 |
| | TTGATAATAT | TCCACAAACA | GCAGAAGAGT | TTAACTTTGC | ACCAAACAAG | ACATTTAAGC | 16560 |
| 25 | TTATCGTGTA | CAGTTTATTA | GCAGCATCAT | TAACTTATGT | TGTCATGATT | TTATACACTG | 16620 |
| | GTTGGTTATC | AACAAGTCAT | CAAAGTTTAA | ATGGGCAGTT | GTGGTTAACA | GGTGCTGLTA | 16680 |
| 30 | CACAAACAGC | ATTTGGTTAT | ATTGGATTAG | GTGTATTAGC | AATTGCAATT | ATGATGGGTA | 16740 |
| | TATTTACTGG | TTTAAATGGA | TTCTTGATGA | GTTCAAGTCG | CTTGTTATTT | TCTATGGGAC | 16800 |
| | GTTCAGGTAT | TATGCCAACA | ATGTTTAGTA | AATTACATAG | TAAATACAAA | ACACCATATG | 16860 |
| 35 . | TCGCAATCAT | ATTCCTAGTA | GGAGTGTCGT | TAATTGCACC | TTGGCTAGGA | AGAACTGCAT | 16920 |
| | TGACTTGGAT | TGTAGATATG | TCATCTACTG | GTGTATCCAT | TGCCTACTTT | ATTACATGTT | 16980 |
| | TGTCTGCAGC | GAAATTATTC | AGTTATAACA | AACAAAGTAA | TACGTATGCA | CCGGTTTACA | 17040 |
| 10 | AAACGTTTGC | TATTATCGGC | TCATTTGTAT | CATTCATTTT | CTTAGCGTTG | TTATTAGTGC | 17100 |
| | CAGGTTCTCC | TGCAGCACTG | ACTGCACCGT | CTTATATTGC | ATTACTTGGA | TGGTTAATCA | 17160 |
| | TCGGTTTAAT | ATTCTTTGTG | ATTCGATATC | CTAAATTGAA | AAATATGGAT | AATGATGAAT | 17220 |
| 15 | TAAGTCGCTT | GATTTTAAAT | AGAAGTGAAA | ATGAAGTTGA | TGATATGATT | GAAGAACCTG | 17280 |
| | AAAAAGAAAA | AACTAAATAA | TAAAAGAATC | GCACAATAAA | CCTTCTTCAT | TCGGAGGCGT | 17340 |
| so | ATCGTGCGAT | TTTTTGTATT | ATAAATTGAC | ATTTAAGACG | AGGCAGCTGA | ACCTTATATA | 17400 |
| - | TAATTGCTAA | GAGTTAGGGC | TGAGCCATTT | CTAACAAATA | TTTATAATCG | TTTAAAAGAT | 17460 |
| | TTCACGAACC | CAGAAACAAT | TAATTTGGAA | ATTTGGTCGG | CARTARTAR | ACCTA ATCCC | 17520 |

| AAGACTAAAT | TTTTTGTAGC | ATCGTATGCT | AAGCCACCAG | GTACTAATGG | AATGATACCC | 17640 |
|------------|------------|------------|------------|------------|------------|-------|
| GTTACCATAA | AAATGATGGC | AGGTTCTTTT | TGTTTACGAG | CCATATAATG | ACTTAACAAG | 17700 |
| CCTAATGCTA | AACTACCAAA | GAAACTAGAG | TATATAGTGT | GCACATTAAA | GCCGTTGAAG | 17760 |
| AATAAGGTGT | AAACCATCCA | TCCACACGTA | CCAACGAAAC | CACATGATAG | ATATAATTTT | 17820 |
| CTAGGTGCAT | CAAAAATGAC | GCAGAA | | | | 17846 |

10

5

- (2) INFORMATION FOR SEQ ID NO: 110:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5544 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

20

25

30

35

40

45

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 110:

60 ATTGACACTT GGTGAAAGTA ATATCGCCGC GCTATTTTGG CAAAATGGAC ACTTAGAACC TGAGTTACAA GATGAACAGC CAATTAATAT ATTAGGATCT GKTCAAATCA ACGAATGGAA 120 TGGTAATCAA TCACCGCAAA TAATTATTCA AGATATTGCG ATGAATGAAC AGCAAATATT 180 AGATTATAGA AGTAAGCGAA AAAGTTTACC TTTTACAGAA AATGATGAAA ATATTGTCGT 240 GCTTATTCAT CCTAAAAGTG ATAAAGTAAA TGCGAATGAA TATTATTATG GTGAAGAAAT 300 TAAACAACAA ACTGATAAAG TAGTATTAAG AGATTTACCA ACGTCAATGG AAGACTTGTC 360 TAATTCCTTG CAACAACTGC AATTTTCTCA ACTTTATATA GTTTTGCAAC ATAATCATTC 420 GATTTACTTC GATGGTATAC CTAATATGGA TATTTTTAAA AAGTGTTATA AAGCATTAAT 480 AACTAAACAA GAAACAAATA TCCAGAAAGA GGGTATGTTA TTGTGTCAAC ATTTAAGTGT 540 GAAACCAGAT ACACTTAAAT TCATGTTGAA AGTTTTCTTA GACTTAAAAT TTGTAACACA 600 AGAAGATGGT TTAATTCGAA TCAATCAACA ACCTGATAAA AGATCGATTG ATTCCAGCAA 660 AGTATATCAA TTAAGACAAC AACGTATGGA TGTTGAAAAG CAATTATTAT ATCAAGATTT 720 TTCAGAAATA AAAAATTGGA TAAAGTCACA ATTGTCGTGA GCAATTTAGG AGGAAATATT 780 AATGGATTTA AAGCAATACG TATCAGAAGT TCAAGATTGG CCGAAACCAG GTGTTAGTTT 840 CAAGGATATT ACTACAATTA TGGATAATGG TGAAGCATAT GGCTATGCAA CAGATAAAAT 900 TGTAGAATAC GCAAAAGACA GAGATGTTGA TATCGTTGTA GGACCTGAAG CGCGTGGCTT 960 TATCATTGGC TGTCCTGTAG CTTATTCAAT GGGGATTGGC TTTGCACCTG TTAGAAAAGA 1020 AGGGAAATTA CCTCGTGMAG TCATTCGTTA TGAGTATGAC CTAGAATATG GTACAAATGT 1080

55

| | ATTAGCTACT | GGTGGTACGA | TTGAAGCAGC | AATAAAATTA | GTTGAAAAAT | TAGGCGGTAT | 1200 |
|----|-------------------|------------|-------------------|---------------|------------|------------|------|
| | CGTAGTAGGT | ATTGCATTTA | TAATTGAATT | GAAATATTTA | AATGGTATTG | AAAAAATTAA | 1260 |
| 5 | AGATTACGAT | GTTATGAGTT | TAATCTCATA | CGACGAATAA | TAAATAATAT | AATTTTATCA | 1320 |
| | AATGAAATCC | TTCATCAAAT | GTATAAGAAC | CAATGACTTA | ATTAAAAAAG | TTGTTTAAGT | 1380 |
| 10 | TTTCTTAACA | TGAGATGTTA | GGATTTTTTA | TTTACTGAAA | ATGTTAGATG | ATTGAGCATT | 1440 |
| , | ATACCTTAAT | AACATCGTTT | ATTTATTTCA | TAAATTGTAG | TATCATAGAA | CTAATATTTA | 1500 |
| | AAAAATGAAA | CAGTAGATTT | AGGTCGAATT | TTTGTAAAAG | TTTTAAAAGT | AGGAATAGTA | 1560 |
| 15 | TACAAATTAA | ACTCGCTCAA | GTAAAATTAA | TATTACGATT | AATGACGACA | GGATAAATAT | 1620 |
| | TTATCGTCGA | CGGACGTATG | ATTGGTGTGG | GACAAATACT | ATTCAACAAG | AGTACCTAAA | 1680 |
| | TCATTGTTTA | AGGCGAAGTA | ataaatatga | ATGGGGTGTA | TCATATAATG | AACAACGAAT | 1740 |
| 20 | ATCCATATAG | TGCAGACGAA | LTCTTCACAA | AGCAAAATCA | TATTTGTCAG | CAGATGAATA | 1800 |
| | TGAGTATGTT | TTAAAAAGCT | ATCATATTGC | TTATGAAGCA | CATAAAGGTC | AGTTCCGAAA | 1860 |
| | AAACGGATTA | CCATACATTA | TGCATCCTAT | ACAAGTTGCA | GGTATTTTAA | CAGAAATGCG | 1920 |
| 25 | ATTAGACGGA | CCGACGATTG | TCGCAGGTTT | TTTGCATGAT | GTAATTGAAG | ATACACCGTA | 1980 |
| | TACATTTGAA | GATGTAAAAG | AAATGTTCAA | TGAAGAAGTT | GCTCGAATTG | TTGATGGTGT | 2040 |
| 30 | GACGAAGCTT | aaaaagtaa | AATACCGCTC | AAAAGAAGAA | CAACAAGCTG | AAAATCATCG | 2100 |
| 30 | CAAGTTATTT | ATTGCGATTG | CCAAAGATGT | ACGCGTAATT | TTGGTGAAAT | TAGCAGACAG | 2160 |
| ` | ATTACATAAT | ATGCGTACCT | TGAAAGCCAT | GCCGCGCGAA | AAACAAATTA | GAATTTCTCG | 2220 |
| 35 | AGAAACATTA | GAAATTTATG | CACCATTAGC | ACATCGTCTT | GGTATTAATA | CAATCAAATG | 2280 |
| | GGAACTAGAA | GATACGGCTC | TTCGTTATAT | TGATAATGTG | CAATATTTTA | GAATAGTCAA | 2340 |
| | TTTAATGAAG | AAGAAACGTA | GTGaACGTGA | AGCGTATATC | GAAACGGCTA | TTGATAGAAT | 2400 |
| 40 | ACGTACTGAA | ATGGACCGAA | TGAATATCGA | AGGCGATATA | AATGGTAGAC | CTAAACATAT | 2460 |
| | TTACAGTATT | TATCGGAAAA | TGATGAAGCA | GAAAAAACAA | TTTGATCAAA | TTTTTGATTT | 2520 |
| | GTTGGCGATA | CGTGTTATTG | TCAATTCTAT | TAATGATTGT | TATGCGATAC | TTGGGTTGGT | 2580 |
| 45 | GCATACGTTA | TGGAAACCGA | TGCCAGGACG | TTTTAAAGAT | TATATTGCAA | TGCCTAAACA | 2640 |
| , | AAATTTGTAT | CAGTCATTGC | ATACTACAGT | AGTAGGCCCA | AATGGAGACC | CGCTCGAAAT | 2700 |
| | CCAAATACGA | ACGTTTGATA | TGCACGAAAT | TGCTGAGCAT | GGTGTTGCAG | CACACTGGGC | 2760 |
| 50 | TTACAAAGAA | GGTAAAAAAG | TAAGTGAAAA | AGATCAAACT | TATCAAAATA | AGTTAAATTG | 2820 |
| | COTTO D D D C D D | mm.comaa | | 1 macmana 1 a | ~~~~ | | |

| | TGAGTTGCCA | TATGGTGCTG | TGCCGATTGA | TTTTGCTTAT | GCGATTCACA | GTGAAGTAGG | 3000 |
|----|----------------|--------------|---------------------|------------|------------|------------|------|
| | TAATAAGATG | ATTGGTGCCA | AGGTGAATGG | CAAAATTGTA | CCAATTGACT | ATATTTTACA | 3060 |
| 5 | AACAGGCGAT | ATTGTTGAAA | TACGTACTAG | TAAACATTCA | TATGGACCAA | GTCGTGATTG | 3120 |
| | GTTGAAAATT | GTTAAATCGT | CTAGTGCCAA | AGGTAAAATT | AAAAGTTTCT | TCAAAAAACA | 3180 |
| | AGATCGTTCA | TCTAATATTG | AAAAAGGCÇG | AATGATGGTT | GAAGCTGAAA | TAAAAGAGCA | 3240 |
| 10 | AGGATTTAGA | GTCGAAGATA | TTTTGACAGA | GAAAAATATT | CAGGTTGTTA | ATGAAAAATA | 3300 |
| | TAACTTTGCA | AATGAAGATG | ATTTATTCGC | AGCTGTAGGA | TTTGGCGGCG | TGACATCCTT | 3360 |
| 15 | ACAGATTGTT | AATAAATTAA | CTGAAAGACA | ACGTATTTTA | GATAAACAAC | GTGCTTTAAA | 3420 |
| | TGAAGCACAA | GAAGTTACGA | AATCATTGCC | TATTAAAGAC | AACATCATTA | CTGATAGTGG | 3480 |
| | TGTCTATGTA | GAAGGTTTAG | AAAATGTACT | TATCAAGTTG | TCAAAATGTT | GTAATCCTAT | 3540 |
| 20 | ACCAGGTGAT | GATATTGTAG | GTTATATCAC | CAAAGGTCAC | GGTATTAAAG | TACATCGCAC | 3600 |
| | TGATTGCCCA | AATATTAAGA | ACGAAACTGA | ACGACTAATT | AATGTTGAAT | GGGTAAAATC | 3660 |
| | AAAAGACGCA | ACTCAAAAAT | ATCAGGTTGA | TTTAGAGGTA | Atgcgtatga | CCGAAATGGC | 3720 |
| 25 | TTGTTGAATG | AAGTACTACA | AGCTGTTAGC | TCGACAGCCG | GCAATTTAAT | TAAAGTTTCA | 3780 |
| | GGACGTTCAG | ATATTGATAA | AAATGCAATA | ATAAATATTA | GTGTCATGGT | GAAAAACGTG | 3840 |
| | AATGATGTTT | ATCGTGTGGT | AGAAAAGATC | AAACAACTTG | GTGATGTTTA | TACAGTAACA | 3900 |
| 30 | AGAGTTTGGA | ACTAGAGGTG | CAAAATATGA | AAGTAGTTGT | ACAAAGAGTT | AAAGAAGCAT | 3960 |
| | CGGTGACGAA | TGATACATTA | AATAATCAAA | TCAAAAAAGG | ATATTGTTTA | TTAGTCGGTA | 4020 |
| 35 | TCGGTCAGAA | CTCTACAGAG | CAAGATGCAG | ATGTAATTGC | AAAGAAAATT | GCTAATGCAA | 4080 |
| | GATTATTTGA | AGATGACAAT | AATAAATTAA | ACTTTAATAT | CCAACAAATG | AATGGTGAAA | 4140 |
| | TACTATCAGT | TTCACAATTT | ACTCTCTATG | CAGATGTAAA | AAAAGGTAAC | CGTCCAGGTT | 4200 |
| 40 | TCTCAAATTC | TAAAAATCCT | GATCaAGCGG | TAAAAATTTA | TGAGTATTTT | AATGCaTGCG | 4260 |
| | CTACGAGCGT | ATGGTCTTAC | TGTGAAAACA | GGTGAATTTG | GAACACACAT | GAATGTTAGC | 4320 |
| | ATAAATAATG | ATGGTCCAGT | CACTATTATT | TATGAAAGTC | AGGACGGCAA | AATTCAATGA | 4380 |
| 45 | AAAAAATAGA | GGCATGGTTA | TCTAAAAAGG | GTCTTAAAAA | TAAACGTACT | CTAATAGTAG | 4440 |
| | TGATTGCCTT | TGTCTTATTT | ATCATCTTTT | TATTTTTATT | GCTGAATAGC | AATAGTGAAG | 4500 |
| | ATAGTGGGAA | CATCACGATA | ACTGAAAATG | CTGAATTACG | TACAGGTCCA | AACGCTGCGT | 4560 |
| 50 | ATCCAGTCAT | ATATAAAGTT | GAAAAAGGTG | ACCATTTTAA | AAAGATTGGT | AAAGTAGGTA | 4620 |
| | A ATTOC ATTOCA | ACTTC A ACAT | 3 C 3 T C C 3 C T 3 | ATCA | TTCC2TT | CCATCCCACA | 4680 |

| | TAGTGCTTGA | TCCTGGTCAT | GGAGGTAGTG | ACCAGGGTGC | TTCAAGCAAT | ACTAAATATA | 4800 |
|----|------------|------------|------------|------------|------------|------------|------|
| | AAAGTTTAGA | AAAAGATTAT | ACGTTGAAAA | CAGCAAAAGA | ATTGCAGCGT | ACTTTAGAAA | 4860 |
| 5 | AAGAAGGCGC | AACTGTTAAG | ATGACAAGAA | CAGACGATAC | ATATGTTTCA | CTAGAAAATC | 4920 |
| | GTGATATCAA | AGGCGATGCC | TATTTGAGTA | TACATAATGA | TGCGTTAGAA | TCATCTAATG | 4980 |
| 10 | CAAATGGAAT | GACAGTTTAT | TGGTATCATG | ATAATCAAAG | AGCTTTAGCA | GATACGTTAG | 5040 |
| | ACGCTACGAT | TCAGAAGAAA | GGTCTACTTT | CTAATCGCGG | TTCAAGACAA | GAAAATTATC | 5100 |
| | AAGTGTTAAG | ACAAACAAAA | GTTCCTGCTG | TTTTATTAGA | ATTAGGTTAT | ATTAGTAACC | 5160 |
| 15 | CAACTGATGA | AACGATGATT | AAAGATCAAT | TACATAGACA | AATTTTAGAA | CAAGCAATTG | 5220 |
| | TTGATGGCCT | TAAAATTTAT | TTTTCTGCGT | AGGGCTTGCA | AAAATATGTG | AAAGTAGTTA | 5280 |
| | TCATTGATAT | TGAATTTTAT | AACTAAAACC | GTTAGTATTC | TTGAAATGGT | AAATGAAATA | 5340 |
| 20 | GGTAGCAATC | TAACTAAGAT | TGTGTAGGAA | TATAATCCAT | AGACTGAAAG | ATTATGCTGA | 5400 |
| | GTAGTTTATA | TACATTGAAC | ACAAGAAGAG | GTGCTTTATG | AAAAGTAAAG | CCGTTAAACG | 5460 |
| | TACGTTAAAC | GTTTTGAGTG | GGTTTATTAA | ATGCACGCTT | ATAAAAAGTA | ATGATGATTA | 5520 |
| 25 | CAATTAGGCA | TGTTTTTTAA | ACCA | | • . | • | 5544 |

(2) INFORMATION FOR SEQ ID NO: 111:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1067 base pairs

(B) TYPE: nucleic acid (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

35

40

45

50

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:

| AAAAGATTGC | TAAATATAAAT | GGCATGTTTA | ATATGTTAGA | ACAACAAATC | ATTCATAGCC | 60 |
|------------|-------------|------------|------------|------------|------------|-----|
| AAGATATGGC | TCATTTTAGA | AGTGAATTTT | TTTACGTCAA | TCATGaGCAT | CGAGAAAACT | 120 |
| ATGAAgCACT | CCTAATTTAT | TACAAAAATA | GTATCGACAA | TCCTATTGTA | GATGGTGCAT | 180 |
| GTTATATTTT | AGCCCTACCT | GAAATTTTCA | ATAGTGTTGA | TGTTTTCGAA | TCAGAGTTAC | 240 |
| CATTTTCATG | GGTATATGAT | GAAAATGGCA | TTACCGAAAC | AATGAAATCA | CTTAGCATTC | 300 |
| CATTACAATA | TTTAGTTGCA | GCAGCTTTAG | AAGTAACTGA | TGTGAATATA | TTTAAGCCTT | 360 |
| CAGGATTTAÇ | AATGGGAATG | AATAATTGGA | ATATTGCTCA | AATGCGAATC | TTTTGGCAAT | 420 |
| ATACAGCAAT | TATTAGAAAA | GAAGCACTAT | AACATTAATA | ATTAATTAGC | TATAAAGATG | 480 |
| ATTCACAACA | ATCATCTTTA | TAGCTTTTTT | ATGTCTAATT | ATTTTTCACC | AAAAT MAAA | E40 |

| AATTTTATGT | TTTCAAAAGT | AAACAATCAA | AAGATGTTAG | AAGATTGCTT | CTATATAAGA | 660 | |
|------------------------------------|------------|------------|------------|------------|------------|------|--|
| AAGAAAGTGT | TTGTAGAAGA | ACAAGGCGTC | CCTGAGGAAA | GTGAAATTGA | TGAATATGAA | 720 | |
| TCTGAATCTA | TTCACCTCAT | TGGATATGAT | AATGGACAGC | CAGTTGCCAC | TGCTCGAATA | 780 | |
| CGCCCTATTA | ATGAAACAAC | TGTCAAAATA | GAACGAGTAG | CTGTGATGAA | ATCACATCGT | 840 | |
| GGACAAGGAA | TGGGTAGAAT | GCTTATGCAA | GCTGTAGAAT | CATTAGCTAA | AGATGAAGGT | 900 | |
| TTTTACGTAG | CTACTATGAA | TGCCCAATGT | CATGCTATCC | CATTTTATGA | AAGTTTAAAC | 960 | |
| TTTAAAATGA | GAGGTAATAT | ATTTCTTGAG | GAAGGCATCG | AGCATATTGA | AATGACAAAA | 1020 | |
| AAGTTAACCT | CGCTTAATTA | AAAAAAGTTG | TATCTATTTT | AGAAACA | | 1067 | |
| (2) THEORYMETON FOR CEO TO NO. 112 | | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 112:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18613 base pairs
- (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 112:

| MAGACGLACG | ATAACAACAA | IACGIGIAGI | GAAAGATTTT | AAICIACAIA | TIMCIGACAA | 90 |
|------------|------------|------------|------------|------------|------------|-----|
| AGAATTCATT | GTATTTGTTG | GACCATCGGG | ATGTGGTAAA | TCAACAACAT | TACGAATGGT | 120 |
| TGCTGGACTA | GAGTCTATCA | CATCTGGAGA | TTTTTATATT | GATGGGGAAC | GCATGAACGA | 180 |
| TGTTGAACCA | AAGAATAGAG | ATATTGCGAT | GGTATTTCAA | AACTATGCAT | TATATCCACA | 240 |
| TATGACTGTT | TTTGAAAATA | TGGCATTTGG | GCTAAAGCTA | CGTAAAGTAA | ATAAAAAGA | 300 |
| GATTGAACAA | AAAGTTAATG | AAGCAGCTGA | AATATTAGGA | TTAACTGAGT | ATCTTGGTCG | 360 |
| тааассаааа | GCGTTATCTG | GCGGACAGCG | TCAACGTGTT | GCTTTGGGCA | GAGCTATTGT | 420 |
| TAGGGATGCG | AAAGTCTTTT | TAATGGATGA | ACCATTATCG | AATCTTGATG | CGAAyTtCGA | 480 |
| GTACAAATGC | GCACAGAAAT | ATTGAAATTA | CATAAGCGAC | TTAATACTAC | GACAATTTAT | 540 |
| GTTACACATG | ATCAAACTGA | AGCATTGACG | ATGGCTAGTC | GAATTGTTGT | TTTGAAAGAT | 600 |
| GGCGACATTA | TGCAAGTCGG | CACACCTAGA | GAAATATATG | ATGCCCCTAA | TTGCATATTT | 660 |
| GTGGCGCAAT | TTATCGGCTC | ACCAGCAATG | AATATGTTGA | ATGCTACAGT | TGAAATGGAC | 720 |
| GGATTGAAGG | TAGGAACACA | CCATTTTAAA | TTACATAATA | AAAAATTTGA | AAAGTTAAAA | 780 |
| GCTGCTGGCT | ACTTAGACAA | GGAAATTATT | TTAGGTATTC | GAGCTGAAGA | CATTCATGAA | 840 |
| CAACCAATAT | TTATTCAAAC | TTCTCCAGAG | ACACAATTTG | AATCTGAAGT | AGTTGTATCC | 900 |

| | AAATTAGATT | CAAGAACTCA | AGTGATGGCG | AACGACAAGA | TTACACTAGC | ATTTGATATG | 1020 |
|----|------------|-----------------|-------------|------------|------------|------------|------|
| | AATAAGTGTC | ACTITITIGA | TGAAAAAACA | GGAAATCGTA | TCGTCTAAGG | GGGAGTATTC | 1080 |
| 5 | ATGTCTAAAA | TTTTAAAATG | TATCACGTTA | GCCGTGGTAA | TGTTATTAAT | CGTAACTGCA | 1140 |
| • | TGTGGCCCTA | ATCGTTCGAA | AGAAGATATT | GATAAAGCAT | TGAATAAAGA | TAATTCTAAA | 1200 |
| 10 | GACAAGCCTA | ACCAACTTAC | GATGTGGGTG | GATGGCGACA | AGCAAATGGC | GTTTTATAAA | 1260 |
| 70 | AAAATTACGG | ATCAATATAC | TAAAAAAAACT | GGCATCAAAG | TAAAGCTTGT | AAATATTGGT | 1320 |
| | CAAAATGATC | AACTAGAAAA | TATTTCGCTA | GACGCTCCTG | CAGGAAAAGG | TCCAGATATC | 1380 |
| 15 | TTTTTCTTAG | CACATGATAA | TACTGGAAGT | GCCTATCTAC | AAGGCTTAGC | TGCTGAAATC | 1440 |
| | AAATTATCAA | AAGATGAGTT | GAAAGGTTTC | AATArGCAAG | CACTTAAAGC | GATGAATTAT | 1500 |
| | GACAATAAGC | AACTAGCATT | GCCAGCTATC | GTTGAAACAA | CCGCACTTTT | TTATAATAAA | 1560 |
| 20 | AAATTAGTGA | AAAATGCACC | GCAAACGTTA | GAAGAAGTTG | AAGCTAATGC | TGCCAAACTA | 1620 |
| | ACTGATAGTA | AAAAGAAACA | ATACGGTATG | TTATTTGATG | CTAAAAATTT | CTATTTTAAT | 1680 |
| | TATCCGTTTT | TATTCGGCAA | TGATGATTAT | ATTTTCAAGA | AAAATGGCAG | TGAATATGAT | 1740 |
| 25 | ATTCATCAGC | TAGGACTAAA | TTCAAAACAT | GTCGTCAAGA | ATGCTGAACG | ATTACAAAAA | 1800 |
| | TGGTACGACA | AAGGGTATCT | TCCTAAGGCA | GCAACACATG | ATGTCATGAT | TGGTCTTTTT | 1860 |
| 30 | AAAGAAGGAA | AAGTAGGACA | ATTTGTCACT | GGACCGTGGA | ACATTAATGA | ATATCAAGAA | 1920 |
| | ACGTTTGGTA | AAGATTTAGG | AGTAACAACA | TTACCTACAG | ATGGTGGCAA | ACCTATGAAA | 1980 |
| | CCATTTCTAG | GTGTACGTGG | TTGGTATTTA | TCTGAATATA | GTAAACATAA | GTATTGGGCT | 2040 |
| 35 | AAAGATTTAA | TGCTGTATAT | CACTAGTAAA | GATACATTAC | AAAAATATAC | AGATGAAATG | 2100 |
| | AGCGAAATTA | CTGGACGTGT | TGACGTGAAA | TCATCTAATC | CAAATTTAAA | AGTGTTTGAA | 2160 |
| | AAGÇAAGCAC | GTCATGCTGA | ACCGATGCCT | AATATTCCTG | AAATGCGACA | AGTTTGGGAA | 2220 |
| 40 | CCGATGGGCA | ATGCAAGCAT | ATTTATTTCA | AATGGTAAGA | ATCCTAAACA | AGCGTTAGAT | 2280 |
| | GAGGCGACGA | ATGATATAAC | GCAAAATATT | AAGATTCTTC | ATCCATCACA | AAATGATAAG | 2340 |
| | AAAGGAGATT | AGTTATGACG | AAACGTAACC | CTAAATTAGC | GGCATTATTA | TCTGTTATAC | 2400 |
| 45 | CTGGTTTGGG | ACAGTTTTAT | AATAAAAGAC | CCATTAAAGG | GACGATATTT | TTTATCTTTT | 2460 |
| | TCATCAGTTT | TATTTCTGTT | TTTTATAGCT | TTTTAAATAT | TGGTTTTTGG | GGATTGTTCA | 2520 |
| 50 | CATTAGGGAC | AGTACCTAAG | TTAGACGATT | CTCGTGTCTT | ACTTGCACAA | GGTATTATTT | 2580 |
| 50 | СТАТСТТАСТ | CGTTGCTTTC | GCAATCATGC | TATATATCAT | TAATATTTTA | GATGCATATC | 2640 |
| | GTAATGCTGA | אמדידאאי | CGCAATGAGG | AAATAAACCA | TOCGAAGOGO | GTATCGTGGC | 2700 |

| | TGTAGTTGTA | TTTCCATTAA | TAYYTATGTT | TGGAGTAGCA | TTTACAAATT | ACAATTTATA | 2820 |
|----|---------------|------------|----------------|-------------------|------------|---------------|------|
| | CAACGCGCCT | CCGAGACACA | CATTAGAATG | GGTTGGTTTA | GATAACTTTA | AAACGTTATT | 2880 |
| 5 | CACAATTGGC | GTTTGGCGTA | AAACATTTTT | CAGTGTTATT | ACTTGGACAT | TAGTATGGAC | 2940 |
| | GCTTGTTGCA | ACGACACTTC | AAATTGCATT | AGGGCTGTTT | TTGGCAATTA | TTGTAAATCA | 3000 |
| | CCCTGTCGTC | AAAGGTAAGA | AATTTATCCG | TACTGTGTTA | ATCCTACCTT | GGGCTGTACC | 3060 |
| 0 | ATCATTTGTG | ACAATTTTAA | TATTTGTAGC | GTTATTTAAT | GATGAATTTG | GTGCGATAAA | 3120 |
| | TAATGATATT | TTGCAACCTT | TATTAGGTGT | AGCACCAGCA | TGGTTAAGTG | ATCCGTTTTG | 3180 |
| 5 | GGCAAAAGTG | GCATTAATCG | GCATTCAAGT | ATGGCTTGGA | TTCCCATTTG | TCTTTGCACT | 3240 |
| | GTTCACTGGA | GTACTGCAAA | GTATTTCATC | AGATTGGTAC | GAAGCAGCAG | ATATGGATGG | 3300 |
| | TGCGTCTAGT | TGGCAAAAGT | TTAGAAACAT | CACATTCCCG | CATGTCATTT | ACGCCACAGC | 3360 |
| ?0 | GCCATTGTTA | ATTATGCAAT | ATGCAGGTAA | TTTCAATAAT | TTTAATCTTA | TTTATCTATT | 3420 |
| | TAATAAAGGC | GGTCCACCAG | TGTCAGGGCA | GAATGCTGGT | AGTACAGATA | TCTTGATATC | 3480 |
| | TTGGGTGTAT | AATCTGACAT | TTGAGTTTAA | CAACTTCAAC | ATGGGTGCAG | TTGTGTCATT | 3540 |
| ?5 | AATTATTGGA | TTTATTGTTG | CTATTGTCGC | ATTTATTCAA | TTCAGACGTA | CAAGTACGTT | 3600 |
| | TAAAGATGAG | GGAGGTTTAT | AAGATGACAA | AGAAGAAAA | CATATTAAAA | GCAATCGGTA | 3660 |
| | TTTACAGTTT | TATAGCGATG | ATGTTTGTCA | TCATTTTATA | TCCACTACTG | TGGACATTTG | 3720 |
| 30 | GCATTTCCCT | TAATCCAGGT | ACGAACTTGT | ATGGTGCCAA | AATGATACCA | GÀCAATGCAA | 3780 |
| | CATTTAAAAA | TTATGCATTC | TTACTATTCG | ATGACAGTAG | TCAATACCTG | ACTTGGTATA | 3840 |
| 35 | AAAATACGCT | TATCGTAGCA | TCTGCAAATG | CACTGTTTAG | TGTGATATTT | GTCACGTTAA | 3900 |
| | CAGCATATGC | TTTTTCTAGA | TATCGCTTTG | TTGGTCGTAA | ATACGGGCTG | ATTACATTTT | 3960 |
| | TGAŤTTTACA | AATGTTCCCT | GTATTAATGG | CAATGGTCGC | AATCTATATT | TTGCTAAATA | 4020 |
| 10 | CAATTGGATT | ATTAGATTCT | TTATTTGGAC | TAACACTGGT | ATATATTGGT | GGATCAATAC | 4080 |
| | CGATGAATGC | CTTTTTAGTG | AAAGGTTACT | TCGATACGAT | TCCAAAAGAA | CTTGATGAAT | 4140 |
| | CTGCCAAAAT | TGATGGTGCA | GGGCATATGC | GTATTTTCTT | ACAAATTATG | CTTCCATTAG | 4200 |
| 45 | CTAAGCCGAT | TTTAGCAGTT | GTTGCTTTGT | TCAATTTTAT | GGGGCCATTT | ATGGACTTTA | 4260 |
| | TATTACCTAA | AATACTATTA | AGAAGTCCTG | AAAAATTCAC | ATTAGCAGTT | GGATTGTTCA | 4320 |
| | ACTITATIAA | TGATAAGTAT | GCAAATAATT | TCACAGTGTT | TGCAGCAGGG | GCAATTATGA | 4380 |
| 50 | TTGCAGTACC | TATAGCAATC | GTATTCTTGT | TCTTGCAACG | CTATTTAGTA | TCAGGTTTAA | 444 |
| | 63.3.63.66EGG | CACAAAACC | TA COTTO A A A | TT 3 C C 3 CT C C | CCCACAATTC | ATA A AGA ACC | 450 |

| | GGGTGTGGTG | GTATTGCGAA | TGGCAAGCAC | ATGCCAAGTT | TACAAAAAGT | TGAAAATGTT | 4620 |
|----|------------|----------------|------------|------------|------------|------------|------|
| • | GAAATGATCG | CATTTTGTGA | CGTAGACATT | TCGAAAGCAG | CGAGTGCGGC | AGAAGCATAC | 4680 |
| 5 | GGAACTGACA | ATGCAAAGGT | TTATGATGAT | TACAAAGCAT | TGTTAAAAGA | TGACACGATT | 4740 |
| | GATGTTATCC | ATGTTTGTAC | GCCAAATGAC | TCGCATTGTG | AAATTACTGT | AGCAGGGTTG | 4800 |
| 10 | CATGCTGGTA | AACATGTGAT | GTGTGAAAAA | CCAATGGCTA | AAACGACAGC | AGAAGCTCAA | 4860 |
| 70 | AAAATGATAG | ATACAGCTAA | ATCAACAGGT | ааааааттаа | CAATAGGTTA | TCAAAATCGT | 4920 |
| | TTCCGAGCAG | ATAGTCAATT | TTTACATCAA | GCAGCGCAAC | GTGGCGACTT | AGGAGACATT | 4980 |
| 15 | TACTTCGGAA | AGGCACATGC | CATTCGTCGT | CGAGCAGTAC | CAACATGGGG | TGTCTTTCTA | 5040 |
| | GACGAAGAAG | CTCAAGGTGG | AGGACCATTA | ATCGATATCG | GTACACACGC | TTTAGATTTA | 5100 |
| | ACGTTATGGA | TGATGGATAA | TTATGAACCA | GAATCAGTGA | TGGGTTCAAC | ATTCCATAAA | 5160 |
| 20 | TTAAATAAAC | AGCATCATGC | GGCAAACGCT | TGGGGTTCAT | GGAATCCAGA | TGAATTTACA | 5220 |
| | GTTGAAGATT | CTGCGTTTGG | ATTTATTAAA | ATGAAGAATG | GAGCGACGAT | CATTTTAGAA | 5280 |
| | TCCGCTTGGG | CGATTAATTC | TTTAGAAGTG | GATGAGGCAA | AATGTTCATT | ATCAGGAACT | 5340 |
| 25 | AAAGCAGGTG | CTGATATGAA | AGATGGTCTA | CGTATTCATG | GTGAAGACAT | GGGTACACTT | 5400 |
| | TATACCAAAC | ACGTTGAATT | GGAAAACAAA | GGCGTCGACT | TTTATGAAGG | TAATGAAGTG | 5460 |
| 20 | GATGAAGCTG | AAGAAGAAGC | AAAAGCTTGG | ATTGATGCAG | TTGTAAATGA | TACTGAACCA | 5520 |
| 30 | GTTGTGAAAC | CGGAACAAGC | AATGGTAGTT | ACAAAAATTC | TTGAAGCGAT | TTATCAGTCT | 5580 |
| | GCAAAATCAG | GCAAAGCAAT | TTACTTTGAA | TAACATCATA | CGGTAAGGAG | GCACATCATG | 5640 |
| 35 | ACAAAATTAA | AAGTTGGTGT | GATAGGTGTT | GGTGGTATTG | CACAAGACCG | TCATATTCCA | 5700 |
| | GCATTGCTGA | AACTCAAAGA | CACAGTCTCA | TTAGTTGCAG | TACAAGATAT | TAATACAGTG | 5760 |
| | CAGATGATTG | ATGTTGCGAA | gCGCTTTAAT | ATACCTCATG | CAGTTGAGAC | ACCTAGCGAG | 5820 |
| 40 | CTGTTTAAAC | TTGTTGATGC | GGTGGTCATT | TGTACACCTA | ATAAATTCCA | TGCTGATCTT | 5880 |
| | TCTATAGAAG | CATTGAACCA | TGGTGTCCAT | GTATTGTGTG | AAAAGCCAAT | GGCGATGACG | 5940 |
| | ACGGAAGAGT | GTGATCGCAT | GATTGAAGCG | GCTAATAAAA | ATCACAAATT | ATTAACTGTC | 6000 |
| 45 | GCATATCATT | ATCGTCACAC | AGATGTGGCA | ATTACTGCTA | AAAAAGCAAT | TGAATCAGGT | 6060 |
| | GTGGTTGGTA | AACCTTTAGT | AGCACGTGTA | CAAGCGATGC | GTAGGCGTAA | AGTGCCTGGC | 6120 |
| 50 | TGGGGTGTTT | TTACCAATAA | AGCGTTGCAA | GGTGGCGGTA | GTTTAATCGA | TTATGGTTGC | 6180 |
| 50 | CACTTGTTAG | ACTTATCTTT | GTGGCTACTA | GGTAAAGATA | TGGTGCCGCA | TGAAGTGCTA | 6240 |
| | CCAAAAACAT | מרבבררבברה מדב | GAGCAAACAA | CCGAATCAAA | TTAATGATTG | GGGAACATTT | 6300 |

| | GCAAGCATGC | AGTTTGAATG | TTCGTGGTCT | GCAAATATCA | AAGAAGATAA | GGTTCACGTT | 6420 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | AGTTTATCAG | GAGAAGATGG | CGGTATCAAT | TTATTTCCAT | TTGAAATATA | TGAGCCCCGC | 6480 |
| 5 | TTTGGAACTA | TTTTTGAAAG | CAAAGCTAAT | GTTGAGCATA | ACGAAGACAT | TGCTGGTGAG | 6540 |
| | AGACAGGCGC | GTAACTTTGT | CAATGCGTGT | TTAGGGATAG | AAGAGATTGT | GGTGAAACCG | 6600 |
| 10 | GAAGAAGCAC | GCAATGTAAA | TGCCCTTATA | GAAGCGATTT | ATCGTAGCGA | TCTTGATAAC | 6660 |
| 10 | AAGAGCATAC | AACTTTAATG | ATTATCATAT | ATGATACAAA | ATTCTCAATA | TAAAAAGAAG | 6720 |
| | GAGTGCTTTT | CAATGAAAAT | AGGTGTATTT | TCAGTATTAT | TTTACGATAA | AAATTTTGAA | 6780 |
| 15 | GATATGTTAG | ATTATGTCTC | AGAATCTGGA | TTGGATATGA | TTGAAGTTGG | AACAGGTGGT | 6840 |
| | AACCCAGGAG | ATAAATTTTG | TAAGTTAGAT | GAGTTGTTAG | AAAATGAAGA | CAAGCGCCAA | 6900 |
| | GCATTTATGA | AGTCAATCAC | AGACAGAGGC | TTACAAATAA | GTGGTTTCAG | TTGTCATAAC | 6960 |
| 20 | AATCCAATTT | CTCCAGATCC | GATAGAAGCG | AAAGAAGCCG | ATGAAACGTT | ACGTAAAACA | 7020 |
| | ATCCGTTTAG | CAAATCTATT | AGACGTGCCA | GTTGTTAATA | ÇATTTTCTGG | CATTGCAGGA | 7080 |
| | TCAGATGATA | CCGCTAAAAA | GCCTAATTGG | CCTGTTACAC | CTTGGCCAAC | AGCCTACTCT | 7140 |
| ?5 | GAAATTTATG | ATTATCAGTG | GAATGAAAAG | TTGATACCAT | ATTGGCAAGA | TTTAGCTGAG | 7200 |
| | TTTGCAAAAG | AGCAAGATGT | AAAAATTGCC | ATAGAGTTGC | ATGCAGGATT | TTTAGTGCAT | 7260 |
| 3 <i>0</i> | ACACCATATA | CAATGTTGAA | GTTACGTGAG | GCTACAAATG | AATATATCGG | TGCTAACTTA | 7320 |
| 30 | GATCCTAGTC | ATCTATGGTG | GCAAGGTATT | GACCCAATTG | CTGCGATTCG | CATATTAGGC | 7380 |
| | CAAGCAAATG | CAATTCATCA | CTTCCATGCT | AAAGATACGT | ATATTAATCA | AGAAAATGTA | 7440 |
| 35 | AATATGTATG | GTCTAACTGA | TATGCAACCA | TATGGTAACG | TTGCGACAAG | AGCATGGACA | 7500 |
| | TTCCGTACAG | TTGGTTATGG | ACATAGTCCA | TATGTATGGG | CAGATATCAT | AAGTCAACTT | 7560 |
| | ATTATTAATG | GATATGATTA | TGTATTAAGT | ATTGAACATG | AAGATCCTAT | TATGTCAGTA | 7620 |
| 40 | GAAGAAGGTT | TCCAAAAAGC | TTGTCAAACT | TTGAAATCTG | TTAATATTTA | CGACAAGCCA | 7680 |
| | GCAGACATGT | GGTGGGCATA | ATACGAACTC | GAGGTTAGTC | TGAAGTTTGT | CTGAAGTAAG | 7740 |
| | ACTGGTGGCA | GTGTTGAATA | AATGCATATG | TCGCCAAGCC | ATTGCCAAAA | ATTTCACACC | 7800 |
| 45 | TTAAATCAAG | TCATTGTTTG | TAAAGAAGGT | GTACTTTATA | TAAGTATATA | GCGATGGTCA | 7860 |
| | TACCCATTCA | CAGTAACAAT | CCTCACCATT | GAAAAGAGTA | TATAACCTTT | TCAATAGTGA | 7920 |
| 50 | GGTATATGAT | AATAAAAAAA | GCCTGTTGTC | ACAATGGTCA | TAGACACGAC | ATACTTTAAA | 7980 |
| 50 | GGTTTCTGAA | TATAATATTT | CAGAATGCAC | TTTAAAGATG | GACGTCGATG | TAGACTAAAG | 8040 |
| | TGATGACAGG | CTTTCATCTT | TTTAAATATT | CATTAATTTC | TCTTCTTGTT | TAATACGTAC | 8100 |

| | TAATACACC | 3 ATTAATTCAG | GAATGATGTT | TAAGAAGTAA | TTTGGGTGTT | TTGTAATTTT | 822 |
|----------|------------|---------------|------------|------------|------------|------------|-------|
| _ | ATATAATCC | A. GATTTAATAA | TAGGATGGTT | AGGTAAAATG | ATTAATTTA | ATGTCCAAAT | 828 |
| . | ACCACCTAA | A GTTTTAATAA | CCATAAATAA | CATGATATAA | GCAAAGATTA | АТАТААСТАА | 834 |
| | GCCAATACCA | 1 TTTGCAAAGC | TAAATGTATC | TTTATTAATA | AATGCCTCTA | CACCAGCCAA | 8400 |
| 10 | TACATAAATT | AAAACGTGTG | TTATTGCTAA | AAACTTCGAA | TTTTTAACGC | CATATTCAAC | 8460 |
| | TGCACCGTCT | GCTTTTAATT | GTTTTGAGTG | ATTAATAGAT | ATCTTTAAGC | TGACAAGTCT | 8520 |
| | GATACAGAAA | AAGATAAGTA | ATATAGATAG | AATCATGATG | TCCTCCGTCA | TTATGTCATA | 8580 |
| 15 | TGTATAAGCG | TTGATTTTGA | CAACATAAAG | TATTTTATAG | ATAAAGCTTG | ТСАААТАСТА | 8640 |
| | TTAACTATTI | ATTAATTTTA | GTACATAAAT | ATGTTTCTAA | GTATGTGTTT | ATGTTCAGTA | 8700 |
| | TTTTGGATAA | TTTAATAATT | TTAAGGATAT | TAAGCGCTTA | CACCGACGTG | ATATATTTGG | 8760 |
| 20 | CTTAACGAAA | ATGATTGAGG | TGACAGAGAT | GAACTTTTTT | GATATCCATA | AGATTCCGAA | 8820 |
| | CAAAGGCATT | CCATTATCGG | TACAACGTAA | ATTATGGCTT | AGAAACTTCA | TGCAAGCTTT | 8880 |
| | CTTCGTAGTG | TTCTTTGTTT | ATATGGCTAT | GTATTTAATT | CGAAACAACT | TTAAGGCGGC | 8940 |
| 25 | ACAACCGTTT | TTAAAAGAGG | AAATTGGATT | ATCTACATTA | GAACTTGGTT | ATATCGGATT | 9.000 |
| | AGCATTTAGT | ATCACGTACG | GTTTAGGAAA | AACATTACTT | GGATATTTTG | TCGATGGACG | 9060 |
| 30 | TAACACAAAA | CGTATTATCT | CGTTCTTACT | TATCTTATCT | GCGATTACAG | TTTTAATTAT | 9120 |
| | GGGATTTGTT | TTAAGTTACT | TTGGTTCTGT | AATGGGATTA | TTAATTGTAC | TTTGGGGACT | 9180 |
| | TAACGGGGTG | TTCCAATCAG | TTGGTGGACC | TGCAAGTTAT | TCAACGATTT | CAAGATGGGC | 9240 |
| 35 | GCCAAGAACG | AAACGTGGCC | GATACTTAGG | ATTCTGGAAT | ACATCACATA | ATATCGGTGG | 9300 |
| | TGCCATAGCA | GGTGGTGTTG | CACTTTGGGG | TGCTAATGTA | TTCTTCCATG | GAAATGTTAT | 9360 |
| | AGGGATGTTC | ATTTTCCCAT | CGGTGATTGC | ATTACTTATT | GGTATCGCAA | CATTATTTAT | 9420 |
| 40 | CGGAAAAGAT | GATCCGGAAG | AATTAGGATG | GAATCGTGCT | GAAGAAATTT | GGGAAGAGCC | 9480 |
| | GGTCGATAAA | GAAAATATTG | ATTCTCAAGG | TATGACGAAA | TGGGAGATCT | TTAAAAAATA | 9540 |
| | TATCCTGGGA | AATCCTGTTA | TATGGATTCT | ATGTGTTTCA | AACGTCTTTG | TATACATTGT | 9600 |
| 15 | ACGAATCGGT | ATTGATAACT | GGGCACCGTT | ATATGTGTCA | GAGCATTTAC | ACTTTAGTAA | 9660 |
| | AGGCGATGCA | GTTAATACGA | TATTCTACTT | TGAAATTGGT | GCATTAGTTG | CAAGTTTATT | 9720 |
| -0 | ATGGGGCTAC | GTATCAGACT | TATTAAAAGG | TCGTCGTGCA | ATTGTAGCTA | TTGGCTGTAT | 9780 |
| 50 | GTTTATGATT | ACATTTGTTG | TCTTATTCTA | CACAAATGCT | ACAAGTGTCA | TGATGGTTAA | 9840 |
| | CATTTCATTG | TTTGCATTAG | GTGCGTTAAT | CTTTGGTCCG | CAATTATTAA | TTGGTGTATC | ganh |

| | CGCGTATCTA | TTCGGTGACT | CAATGGCGAA | AGTTGGTTTG | GCGGCTATTG | CTGATCCAAC | 10020 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | ACGTAACGGT | TTAAACATCT | TTGGATATAC | ATTAAGTGGA | TGGACAGATG | TTTTCATCGT | 10080 |
| 5 | CITCTATGTT | GCATTATTCC | TAGGCATGAT | TCTATTAGGA | ATCGTTGCTT | TCTATGAAGA | 10140 |
| | AAAGAAAATT | AGAAGTTTAA | AAATTTAATA | TAAATCGGAT | TAAAAGTATC | GCCAATCTAT | 10200 |
| | TGCAATATAG | TTGGCAATCC | TGCCCCGACG | GCATGTGCGT | GAAGAGATGA | AAGATACTGC | 10260 |
| 10 | TTCTACCCTT | GCAAATATAT | CATCTCTATG | TCTCGGGGCA | GATCATAATT | CCCTGTTATG | 10320 |
| | AAGTATCCTT | ATTTGCCCGA | CTTAGGGTGA | CTCAATGAAT | TTACTCCTTA | CAATAAAGAC | 10380 |
| 15 | ATATAGCGGT | GTCAATATTG | TAGGGAGTAT | TGTTTTATAT | TTAAACTCTC | TAAAAAGCGG | 10440 |
| | ACTGAAAGAA | AAGTGAAAAC | TTCTCTATCA | GTCCGCTTTT | TCATAGAACA | AAATGGAGGC | 10500 |
| • | GCCATAATCA | TTAGTTATGT | GCTAATCTAT | TTTGCTTGCT | TACAATAATC | ACTTGGCGAC | 10560 |
| 20 | ATTTGTAAAT | ATTTTTTAAA | ATGATAGCTA | AACATTTTAT | ACTCTGAAAA | GCCTACTTTG | 10620 |
| | TCTGCAATTT | CATAGTGTTT | GTAATGTCGA | TCTAACAATT | GCAGAGATTG | TAAAATACGA | 10680 |
| | TAGCGATTTA | AATAATCGAC | AATTGTAATA | CCAACATGAT | CTTTAAATGT | TCGCATCGCA | 10740 |
| 25 | TACGATTCAC | TAACATCGAT | ATGTTGAATT | AAATCTGAAA | CAGECACTTT | CGTTTGATAA | 10800 |
| | GATTGCTTAA | TTTGATCCAC | AATCTGGTTT | ACATAATAAT | CATCGTATTC | TACTTTTAAT | 10860 |
| | AGTGGTTGGA | AGGCATCATG | ACAAGATGCT | AAGCTACGGC | CGTTCTGTGA | TTGTTGCTCT | 10920 |
| 30 | AATAAGGTAC | GGACAAGTCT | TCCTAAAATA | ACTTCTAATT | GTGCATGGTC | TACTGGTTTT | 10980 |
| | AATAAATAAT | CAAGAACATG | ATGTTGAATG | CCGGCTTTCA | TATATTCAAA | GTCATCGTAA | 11040 |
| 35 | CTCGATAATA | TGATGACATT | ACAATCTAGA | TGCGCAATAT | CATTGAGTAA | ATCGACGCCA | 11100 |
| | TTTTTACGTG | GCATACGAAT | ATCAGTAATT | ACTAATTCTG | GCTGATGTTG | TTGAATTAGT | 11160 |
| | GATAATGCTT | CAACACCATC | TTTAGCAGTG | TATATTGTAT | TGAAATGATA | GTCTCCCCAA | 11220 |
| 40 | GGAATGATTT | GCTTTAATCC | TTCTCGAATA | ATTCGTTCAT | CATCACAAAT | AACTACCTTA | 11280 |
| | AACATCTACA | TTCCCCCTTG | AAAGTGGTAT | TTTATAACAA | ATTAACGTAC | CTTGATTACG | 11340 |
| | CTTTGAAAAA | ATATGGAGTC | GTGCATGTGA | ACCATATTGA | ATCATTGCTT | TATTGTGTAA | 11400 |
| 45 | ATGATTTAAT | CCCAAATGCT | TAGTATCAAA | TACATCATTA | TTAAGAGATT | GGCGTACATA | 11460 |
| | TTGCAGGCGA | GATGACGACA | TCCCGATACC | ATTGTCGCAA | ACTAAAACAT | GTAAATTCTG | 11520 |
| | ACGTGCCAAT | GTCAGGCGTA | TAGTAATGTC | CAATGACTCA | GTATCTCTAC | CATGTTTAAT | 11580 |
| 50 | AGCATTTTCT | ATGAGTGGCT | GAAGCATCAT | TTTACCAATT | GTCTGGTGAC | GCGCTTCTTC | 11640 |
| | AGAACTTTCA | ATATGGAGCT | TAATCATGTC | ATCAAAACGG | aTGTTTTGTA | TTGCAACATA | 1170 |

| | GTAACGTAAC | ATTTGCGATA | ATTGTTGGAC | CACAGTTtGT | GCTAATTTCG | GAGATAACGT | 11820 |
|-----|--------------------------|------------|-------------|--------------------------|--------------|----------------|-------|
| | AATTAAATAT | TGTATTGTTT | GCATCGTATT | GAATAGGAAA | TGAGGCTGGA | ATTGGCGTTC | 11880 |
| 5 | TATTTCCTTT | AACTGAATAT | CACGCAAGCG | ACGTTCTGTA | TGCTCGATAG | AATGGATCAG | 11940 |
| | TTGCTCATTT | GATTCAAATA | AATCGTAAAT | ATAATTATTA | ATTTCTTCTA | GTTCACTGTT | 12000 |
| | GTTTTTTAAA | GGCGTATATG | TACCTAGATG | ACGATTTTTG | GCATAGTAAA | TTTTTTGAAT | 12060 |
| 10 | AATCGTTTCG | ATATCTTTTG | TTTGTCGTTT | AGCCATATTA | TCTGCGCTAA | TGAAACCAAA | 12120 |
| | TATTACTAGT | AAAACAAGAA | CTACGGCCAT | AACAATTAAC | AACGTGATAC | CATCTTCAAT | 12180 |
| 15 | GTTTTCATGT | ATATCTTTAT | AAATAATGAG | ACGATGGTCA | GCATGGTTTA | ATTTTACAGA | 12240 |
| | TTCATTCATA | AATCCGAATT | GTTGTGGTcT | ATACTTTTCA | CCTATAGTAA | AACGGTCATC | 12300 |
| | GTTGGCGTAT | AAAATATTGT | CATATTGATC | AmCGATAAGT | GCGAATTGTC | GGTTATCTTT | 12360 |
| 20 | CETAATTTCA | CTTAAACGTG | GGGTGTtAGC | CATATAAATt | TTAAGCATAT | ATGTACTATT | 12420 |
| • • | TTTGAATTTA | AGCTGATGCG | TTGAAAATAA | ATACATATTT | TTAGTGTTTA | AATGTTCATA | 12480 |
| | ATTATTGGTT | ATAAACTGAT | TTGGTCCAGA | TAATTCATAA | TAAAGTGTTG | CGGGCTGTTG | 12540 |
| 25 | GKGTATTAAT | TTTAATAATT | CACGTTTTGT | AGCGGTCACA | TCATGATGAT | TTGYTAAATC | 12600 |
| | GAGCTCTTGA | AACGAATTAT | TATGCTGTGT | AATAAATGTC | TGAATCTGCT | TTTCAGTATG | 12660 |
| | ATGTAAAGAT | GACTGACTTT | CATCAACATG | TTGATGAATC | GTACGATGCT | CAATCCAAAT | 12720 |
| 30 | ATAGATGGCA | TAGAAGCTTA | CTAGTCCAAT | AATAATGACT | AAAAATACTG | GAAAAATAGT | 12780 |
| | AGACNCAAAT | AACGATCGTC | TTAATTGATG | TCTATAAGGT | TTGTATGCCn | TCATTGAATC | 12840 |
| 35 | ATCTCCAAAA | ATTTATGATG | TGGAATATCC | GGTAATTTAG | ATTTCGGTAT | TAAAGGTATG | 12900 |
| | TTCTTAAGAT | TTTCGATAGA | CTGATCGCTT | TGTTCACTAA | CATCCTTTCG | AATTGACTTG | 12960 |
| | GCATCGAACT | CTGCAACTAA | TCGTtGTTGT | ACTGAGCGGC | TTGTTAAATA | TTGCACTAAC | 13020 |
| 40 | TTTTTACGCT | TAGGATGAGG | GTGTGCATTT | TTAACTAAAG | CAATrCCATC | AACATTTAAC | 13080 |
| | ATTGTTCCTT | CAATTGGATA | AACGATTGAT | ACAGGATAAC | CTTTGTTTTT | CCATGTGCGT | 13140 |
| | GCATCTTGTT | CGTAGCTTAG | ACCTGCGTAA | TATTTACCTT | TTGCAACATC | TTCAATGACT | 13200 |
| 45 | TTAGACGTCT | TTGACAGTTG | CATCGCATGG | TTTTGGAATT | GATGCACATC | ACTTACTCGA | 13260 |
| 4 | TGATGCATGC | TATAAATAGC | ACGCATATGT | TGATAGCCTG | TCGTTGTTGT | ATTTGGATTT | 13320 |
| | GAGTACGCAA | TTTTACCTTT | AAGTATAGGT | TGTAATAAAT | CTTGATAACC | TCGAATCTTA | 13380 |
| 50 | ATATCTCCTT | GTAAATCTGA | ATTCACTACT | ATAACTGTTG | GCATTAATAG | AAAACTAGTA | 13440 |
| | ה מידיית מידים בי | TGTTCGAGCC | ата атсетет | א אייירירירייריירייריירי | TOTA CA CATO | ATTCCTTC ATTAC | 12500 |

| | CCACGCTCCG | AAAAATCTTC | GTTATGCAAG | TTTGAAAGCA | GTACTTGAGT | AGATCCGTGT | 13620 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | TTAATTTCAA | TTTTGACATG | CTCTTGTTTT | TCAAATTCAT | TTAAAATTGG | ACGAATCAAG | 13680 |
| 5 | TTTGATTGAT | ACGGAGAATA | AACTGTTAAT | ACATTTTTAT | CGGATTCAGA | GTGACGCGTA | 13740 |
| | TTAGCGCATG | CTGaTAAAAA | AATGAGAAAT | AATAGCAAGA | TATAAATTTT | TGATTTCATG | 13800 |
| o | ATATCCCATC | AATTCTATGT | ATATTTTAAT | ACAATAATTT | TAGCAATAAA | TGACGCATAA | 13860 |
| • | GTÄATGTTAA | ATATTTAGAA | ATGTTTATAG | ATGACTTGTT | AAGACGTTGC | AAATGTTGTG | 13920 |
| · | ATAGCACAAA | ATTTTTGTTT | GTCAAGACGA | TTTACCGAGG | CTGTAAAATC | AAACTGTTAT | 13980 |
| 5 | ATTTTATTTG | TAGCTGTTAT | ATAAAAATCG | GCAAGATATT | GAACGGTTCA | AAAGTGAATT | 14040 |
| | TTTACGTCAA | TAAAAGTATT | TAATCCAGTC | TCTTCATATA | TAAAAGTAAA | TCTTTCTAAG | 14100 |
| | TGTTGATTTA | ACGCTTATCA | ACAATCATTT | TTTATAAACA | AATATATACT | CCTAAATTAA | 14160 |
| 0 | CTTTTAAAGC | AATGAAAATA | GTGAACATTA | TAACTGTTGT | GTAACAGAAT | GCAATTAGCA | 14220 |
| | TATTACTGTT | ACACAAATTA | GTACAGTTTC | TATGTTTTGA | CATACATTTG | ATGAAAATTG | 14280 |
| | TACATAATTT | ATGTGAAAAA | AATCACAACA | AACATGCTAC | AATGACTATG | AAAACGTTAA | 14340 |
| 5 | CATAGCATTT | CAAATTCACA | ACATTATACA | GATGGAGGCG | TTTAGTATGT | TAGAAACAAA | 14400 |
| | Taaaaatcat | GCAACAGCTT | GGCAAGGATT | TAAAAATGGA | AGATGGAACA | GACACGTAGA | 14460 |
| 10 | TGTAAGAGAG | TTTATCCAAT | TAAACTACAC | TCTTTATGAA | GGTAATGATT | CATTTTTAGC | 14520 |
| | AGGACCAACA | GAAGCAACTT | CTAAACTTTG | GGAACAAGTA | ATGCAGTTAT | CGAAAGAAGA | 14580 |
| | ACGTGAACGT | GGCGGCATGT | GGGATATGGA | CACGAAAGTA | GCTTCAACAA | TCACATCTCA | 14640 |
| 15 | TGATGCTGGT | TATTTAGACA | AAGATTTAGA | AACAATTGTA | GGTGTACAAA | CTGAAAAGCC | 14700 |
| | ATTCAAACGT | TCAATGCAAC | CATTCGGTGG | TATTCGTATG | GCGAAAgcAG | CTTGTGAAGC | 14760 |
| | TTAÇGGTTAC | GAATTAGACG | AAGAAACTGA | AAAAATCTTT | ACAGATTATC | GTAAAACACA | 14820 |
| 10 | TAACCAAGGT | GTATTCGATG | CATATTCTAG | AGAAATGTTG | AACTGCCGTA | AAGCAGGTGT | 14880 |
| | AATCACTGGT | TTACCTGATG | CATACGGACG | TGGACGTATT | ATCGGTGACT | ATCGTCGTGT | 14940 |
| | AGCTTTATAT | GGTGTAGATT | TCTTAATGGA | AGAAAAAATG | CACGACTTCA | ACACGATGTC | 15000 |
| 15 | TACAGAAATG | TCAGAAGATG | TAATTCGTTT | ACGTGaAGAA | TTATCAGAAC | AATATCGTGC | 15060 |
| | ATTAAAAGAA | TTAAAAGAAC | TTGGACAAAA | ATATGGTTTC | GATTTAAGCC | GTCCAGCAGA | 15120 |
| -0 | AAACTTCAAA | GAAGCAGTTC | AATGGTTATA | CTTAGCATAC | CTTGCTGCAA | TTAAAGAACA | 15180 |
| 50 | AAACGGTGCA | GCAATGAGTT | TAGGTCGTAC | ATCAACATTC | TTAGATATCT | ATGCTGAACG | 15240 |
| | TCACCTTAAA | GCAGGCGTTA | ттастсавас | CGAAGTTCAA | GAAATTATTG | ACCACTTCAT | 15300 |

| | AGACCCAACT | TGGGTAACTG | AATCTATCGG | TGGTGTAGGT | ATTGACGGAC | GTCCACTTGT | 15420 |
|---|------------|------------|------------|------------|------------|------------|-------|
| | TACGAAAAAC | TCATTCCGTT | TCTTACACTC | ATTAGATAAC | TTAGGTCCAG | CTCCAGAACC | 15480 |
| 5 | AAACTTAACA | GTATTATGGT | CAGTACGTTT | ACCTGACAAC | TTCAAAACAT | ACTGTGCAAA | 15540 |
| | AATGAGTATT | AAAACAAGTT | CTATCCAATA | TGAAAATGAT | GACATTATGC | GTGAAAGCTA | 15600 |
| | TGGCGATGAC | TATGGTATCG | CATGTTGTGT | ATCAGCGATG | ACAATTGGTA | AACAAATGCA | 15660 |
| 10 | ATTCTTCGGT | GCACGTGCGA | ACTTAGCTAA | AACATTACTT | TACGCTATCA | ATGGTGGTAA | 15720 |
| | AGATGAAAAA | TCTGGTGCAC | AAGTTGGTCC | AAACTTCGAA | GGTATTAACA | GCGAAGTATT | 15780 |
| 15 | AGAATATGAC | GAAGTATTCA | AGAAATTTGA | TCAAATGATG | GATTGGCTAG | CAGGTGTTTA | 15840 |
| 13 | CATTAACTCA | TTAAATGTTA | TTCACTACAT | GCACGATAAA | TACAGCTATG | AACGTATTGA | 15900 |
| | AATGGCATTA | CATGATACAG | AAATTGTACG | TACAATGGCA | ACAGGTATCG | CTGGTTTATC | 15960 |
| 20 | AGTAGCAGCT | GACTCATTAT | CTGCAATTAA | ATATGCACAA | GTTAAACCAA | TTCGTAACGA | 16020 |
| | AGAAGGTCTT | GTAGTAGACT | TTGAAATCGA | AGGCGACTTC | CCTAAATACG | GTAACAATGA | 16080 |
| | CGACCGTGTA | GATGATATTG | CAGTTGATTT | AGTAGAACGC | TTCATGACTA | AATTACGTAG | 16140 |
| 25 | TCATAAAACA | TATCGTGATT | CAGAACATAC | AATGAGTGTA | TTAACARTTA | CTTCAAACGT | 16200 |
| | TGTATACGGT | AAGAAAACTG | GTAACACACC | AGACGGACGT | AAAGCTGGCG | AACCATTTGC | 16260 |
| | TCCAGGTGCA | AACCCAATGC | ATGGCCGTGA | CCAAAAAGGT | GCATTATCTT | CATTAAGTTC | 16320 |
| 30 | TGTAGCTAAG | ATCCCTTACG | ATTGCTGTAA | AGATGGTATT | TCAAATACAT | TCAGTATCGT | 16380 |
| | ACCAAAATCA | TTAGGTAAAG | AACCAGAAGA | TCAAAACCGT | AACTTAACTA | GTATGTTAGA | 16440 |
| 35 | TGGTTACGCA | ATGCAATGTG | GTCACCACTT | AAATATTAAC | GTATTTAACC | GTGAAACATT | 16500 |
| 33 | AATAGATGCA | ATGGAACATC | CAGAAGAATA | TCCACAGTTA | ACAATCCGTG | TATCTGGTTA | 16560 |
| | CGCTGTTAAC | TTCATTAAAT | TAACACGTGA | ACAACAATTA | GATGTAATTT | CTCGTACATT | 16620 |
| 40 | CCATGAAAGT | ATGTAACAAA | ATTTAAGGTG | GGAGCACTAT | GCTTAAGGGA | CACTTACATT | 16680 |
| on one - product models production - or | CTGTCGAAAG | TTTAGGTACT | GTCGATGGAC | CGGGATTAAG | ATATATATTA | TTTACACAAG | 16740 |
| | GATGCTTACT | TAGATGCTTG | TATTGCCACA | ATCCAGATAC | TTGGAAAATT | AGTGAGCCAT | 16800 |
| 45 | CAAGAGAAGT | CACAGTTGAT | GAAATGGTGA | ATGAAATATT | ACCATACAAA | CCATACTTTG | 16860 |
| | ATGCATCGGG | TGGCGGTGTA | ACAGTCAGTG | GTGGCGAACC | ATTGTTACAA | ATGCCATTCT | 16920 |
| | TAGAAAAATT | ATTTGCAGAA | TTAAAAGAAA | ATGGTGTGCA | CACTTGCTTA | GACACATCGG | 16980 |
| 50 | CTGGATGTGC | TAATGATACA | AAAGCATTTC | AAAGGCATTT | TGAAGAATTA | CAAAAACATA | 17040 |
| | CAGACTTGAT | ATTATTAGAT | ATAAAACATA | TTGATAATGA | CAAACATATT | AGATTGACAG | 17100 |

| | TATGGATTCG | ACATGTCCTT | GTGCCTGGTT | ATTCTGATGA | TAAAGACGAT | TTAATTAAAC | 17220 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | TAGGGGAATT | TATTAATTCT | CTTGATAACG | TCGAAAAGTT | TGAAATTCTG | CCATATCATC | 17280 |
| 5 | AGTTAGGTGT | TCATAAGTGG | AAAACATTGG | GCATTGCATA | TGAATTAGAA | GATGTCGAAG | 17340 |
| | CGCCCGATGA | TGAAGCTGTT | AAAGCAGCCT | ACCGTTATGT | TAACTTCAAA | GGGAAAATTC | 17400 |
| | CCGTTGAATT | ATAAATACAA | TTCAGACCGA | AAAGAAAGCA | TATGCAACTT | CAAGAGTGAA | 17460 |
| 10 | GGGGCATATG | CTTCTTTTTC | AATTGAGTAT | TGAGTATTAG | CAAGACGTAG | TAAGTATATG | 17520 |
| | AGACAACTTC | TACAATGGTT | GAAGGAAGAC | GTTTTTGTAA | GTAGCTATGC | TGATAAAGAA | 17580 |
| 15 | TGTGATGTCT | TGTTAAAGGT | GGGGTTCCAA | TATCATCATT | TAGCTGATGT | TGAATGGGTT | 17640 |
| | ATTATTTGCT | ACTTGCATAT | GAATATGAGT | CTTTTCAAAT | TTTTATTGAC | CCTGAGTAAT | 17700 |
| | GAAAAATATT | AAGATGAAAC | TTAATATTAA | AgCAATGCGG | AGCGTGATTA | TGAAGAGAAT | 17760 |
| 20 | TAGTAAAGAT | ATATGGGCAG | TATTTAAATT | ACTGTATCAA | AATAAAGGGC | GTTTTAGCAT | 17820 |
| | TAATGCCTTA | CTATTGCAGT | TAATCATGAT | TTTTATTAGT | AGTACATACT | TAATTTTACT | 17880 |
| | ATTTAATATG | ATGTTAAAAG | TAGCTGGCAA | AGCCAACTTA | CGATTAACAA | TTGGACGGAA | 17940 |
| 25 | ATCGTTAGTC | ATCCCGCCAG | TGTGATACTT | CTTATTATAT | TCATATTAAG | TGTTGCCTTT | 18000 |
| | CTGATTTATG | TAGAGTTTTC | ATTGTTAGTT | TATATGGTTT | ATGCCGGCTT | TGATCGACAG | 18060 |
| | ATTATTACAT | TTAAATCCAT | TTTTAAAAAT | GCCTTTGTAA | ATGTGCGTAA | ACTCATAGGT | 18120 |
| 30 | GTACCAGTTA | TTTTCTTTGT | CATTTATTTA | ATGTTAATGA | TACCCATTGC | CAACCTAGGA | 18180 |
| | CTAAGTTCAG | TATTAACAAA | AAATATTTAC | ATACCTAAAT | TTTTAACGGA | AGAACTTATG | 18240 |
| 35 | AAAACGACGA | AAGGTATAAT | CATTTACGGT | ACCTTTATGA | TTGCTGTATT | TATATTAAAT | 18300 |
| | TTTAAATTAA | TATTTACTCT | ACCGTTAACG | ATTTTAAACC | GCCAGTCGTT | TAAAAAT | 18360 |
| | ATGĄGACTAA | GTTGGCAAAT | TACGAAGCGA | AATAAGTTTC | GGCTTGTTAT | AGAAATAGTT | 18420 |
| 40 | ATATTAGAAC | TCATCATTGG | TGCGATTTTA | ACATTAATTA | TTTCAGGAGC | AACATATCTT | 18480 |
| | GCTATTTGTG | TAGATGAAGA | AGGAGATAAG | TTTTTAGTCT | CATCAATTTT | ATTTGTTGTA | 18540 |
| | TTGAAAAGCG | CATTGTTCTT | CTATTATKTA | TTtACGAAAT | TATCATTAAT | CAGTGTGTTA | 18600 |
| 45 | GTACTGCACT | TAA | | | | | 18613 |

6.7

....

(2) INFORMATION FOR SEQ ID NO: 113:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1214 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 113:

| | AAAGIIIIAA AAGGGGIGAG ATACTIGGCG AATAATCCAT TCCAGCTTTG CGTTTAAAAG | 60 |
|----|---|------|
| 5 | GAATTATACT TGCCATTGTC GGTGCTTGTT TATGGGGATT AGGTGGTACT GTTTCTGATT | 120 |
| | TCTTGTTCAA ATATAAGAAT ATTAATGTCG ATTGGTACGT CACTGCTCGA CTTGTAGTCA | 180 |
| 10 | GTGGTGTTTT CTTACTTATT ATGTACAAAA TGATGCAACC CAAACGTTCA ATATTTAGCG | 240 |
| 10 | TATTCCAAGA TCGACGTATG TTAGGCAAAT TACTTATCTT CAGTATACTG GGCATGTTAG | 300 |
| | TAGTACAATA TGCTTATATG GCATCTATTA ATACAGGTAA TGCTGCGATT GCAACATTAC | 360 |
| 15 | TACAATACAT TGCGCCAGTT TATATTATTA TTTGGTTTGT CATAAGAGGC GTTGCAAAAC | 420 |
| | TAACATTATT TGATGTGCTT GCTATTATCA TGACACTATT AGGAACATTT TTATTATTAA | 480 |
| | CAAATGGTTC ATTTTCTAAT TTAGTCGTCA ATCCTGCAAG TTTATTCTGG GGTATTTTAG | 540 |
| 20 | CTGGTGTAGC ACTCGCTTTT TACACAATTT ATCCTTCAGA CCTACTTAAC CGCTTCGGTT | 600 |
| | CGATTCTAAT TGTCGGGTGG GCAATGCTTA TTTCTGGTGT TGCGATGAAT TTACGCCATC | 660 |
| | CAATTTGGCA CATTGATATC ACTAAATGGG ACATATCAAT TATATTATTT TTAATCTTTG | 720 |
| 25 | GTATTATCGG TGGTACCGCA CTCGCATTTT ATTTCTTTAT CGACAGTTTA CAATACATAT | 780 |
| | CAGCGAAAGA AACAACATTA TTCGGAACTG TTGAACCTGT CGTAGCCGTT ATCGCAAGCA | 840 |
| | GTCTATGGTT ACATGTGGCA TTCAAACCAT TTCAAATCGT AGGCATCATT CTTATTATGA | 900 |
| 30 | TTTTAATTTT ATTACTATCA CTTAAAAGAC AACCTGAAAC ATTAGATGAA TAAGAAAACT | 960 |
| | CTGATAATCA CTTTAGCAAG TAACTATTAT TTAACAACGT AGTTACCTTA TAGGTGATAT | 1020 |
| 35 | CAGAGTTTTT TATTTTAGTT AATAATATTT TTCACTTGGT ATAAAAAAGC GTCGTCGCTC | 1080 |
| | TGGTAATCGG AAATACTGGA ATAAAATATG GAATTGGGTA ATAATCCCAG GTANTAAAAG | 1140 |
| | TCCATGTTCC GATANCCTNT CCGCANCTCC AACCAAATTT GCCGATAAGG TTCCAAAAGG | 1200 |
| 10 | CATCCTGGGG - GTAC | 1214 |
| ~ | (2) INFORMATION FOR SEQ ID NO: 114: | • |
| | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 9458 base pairs | |
| 15 | (B) TYPE: nucleic acid | • |

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 114: ATTTTGGTTT CATTCACGAT GGGGTNATAC AGCAAACACA NCTAAAATAA CTATCAATAG

(C) STRANDEDNESS: double (D) TOPOLOGY: linear

55

| | CTTAGACAAT | AAAAAATATG | CCACTACAAT | CGCTAATATT | ACGATTAAAA | AAGAAGCGTT | 180 |
|----|------------|------------|------------|------------|------------|------------|------|
| | AACGATTACT | TTCATCGTTG | TTCTATCTCT | GAACATCATA | TTAAAGACAA | CTAGACTAAT | 240 |
| 5 | TGATAATGAA | ACAGCAAAAA | AAGTAATAGC | TAACACTAAT | TTCATCATAA | ATAGACAGAC | 300 |
| | TAAACCTATG | ACTAATAATG | TATTAGAAAT | TACAGCTGAC | GTTTTTAACA | TTCTCGaATT | 360 |
| | AATÄTGCACT | CACCCTTTTT | ATTTAAATAA | CTTACATAAT | CATAATAATA | CATGATGTTT | 420 |
| 0 | CATAGGCCTG | TCGATGATTG | ATTCACAATA | GCACGTGATT | TTTTTGTTTT | TCAATATTAT | 480 |
| | TCATTTATTC | CATCAAAAAC | ACCCTTTTTA | ATTTTTACAA | AAATTAAAA | AAGTGCTCCT | 540 |
| 5 | ACACTGCTTG | CATGTAGAAA | CACTTTTTCA | TTGTAATGTT | ATTCTTCTCG | AGACATACCT | 600 |
| | TTTAGCATAT | TAAGCATGTA | TGTTAAACTA | CGGTTCATGT | CGTCATCTTT | CAATACGCCC | 660 |
| | AATAGACTTC | TTATAGTTGT | CTTAGCATTT | GGACTCGCTT | GATTGGCAAC | GTGTAATCCT | 720 |
| 20 | TTATTAACTT | TATTTAGGAA | GTCGCTTAAA | TCTGATACAT | TGAGTTCACC | TAAAAAAT | 780 |
| • | ACCATTGAAG | CCATATTAGA | TAATAGCCCT | GTATAAATAT | CTTTATTAAG | TTCAACTGCA | 840 |
| | AATTTATTTA | TGATGACTTG | ACGTCCTCGA | ATTGCACCAT | TTAAAGCATC | TAATAGTTTT | 900 |
| ?5 | GCATCATCTA | ATGTTTTAAT | AAGCTTGATT | GCTTTTAATA | TACTATCTTT | ATTCGCTGCA | 960 |
| | ATTGCCTCTG | TAACTTCATT | TAAACTTTCT | AACTTAATTT | GTTCTTCTGA | TTTTTCTAAG | 1020 |
| | CGTCTAATTT | TAGAAGATAT | TCTCTCAGCC | ATTATTTATC | CACCTGATTT | CCCGGGAAAA | 1080 |
| 30 | CATAATCTGA | ACGTTCCCAT | TTTTTCTGTA | CTTGAACACT | GTACTGCGGT | TGACGTTTTT | 1140 |
| | TATTGACACG | GAAATTATTA | GGGTTCAACG | GTGACTTACC | ACGTTTCGTA | ATTACCTCCA | 1200 |
| 35 | AACGACAGCT | AGTACGTTTA | TAAGATGGTG | TATCCGTGTA | TTGATCAACA | TCACTATTAG | 1260 |
| | TTAATAAGTT | AATTGCACCT | AGATCTCCAT | TTTCCATCGC | aTCaTTATTT | AATGGAATAT | 1320 |
| | AGAIITCTTT | ACCTTTAACA | CGATCTGTCA | CGTGAACTTG | TAATACCGCT | TCTCCTGTYT | 1380 |
| 10 | CAGAAATCAG | CTTAACTTCT | GCACCTTCAT | GAATGCCTCT | ATCTTCAGCA | AGCTCTGGAG | 1440 |
| | AAATTTCAAC | AAATGCACGT | GGCACTTTGT | ATTTAATCAT | TGGTGTTTGA | TAAGTCATAT | 1500 |
| | TACCTTCATG | GAAGTGCTCT | AACAATCGAC | CATTGTTTAC | ATGAATATCA | TAAATTTCAT | 1560 |
| 15 | CTTGCTTAAA | GTAATTATCA | AATGATAATG | GGAATAATTT | TGCTTTACCA | TTATCAAAAT | 162 |
| | TGAATCCTTC | TAAGTATAGA | ATAGGCTCAT | CAGTACCATC | AGGTTGTACT | GGCCATTGTA | 1680 |
| | AACTATTGAA | TCCTTCTAAA | CGATCATAAC | TTACCCCAGC | ATATAGAGGT | GTTAAGCGTG | 174 |
| 50 | CTACTTCATC | CATAATTTCA | CTAGGATGCT | TGTAATTCCA | ATCAAATCCT | AATCTATTAG | 180 |
| | CAATTGCTTG | GAAAATTTTC | CAGTCAGGTT | TTRAATCACC | AAGAGGTTCT | AATGCTTGGT | 186 |

| | TTGCTGGCAA | TACAACATCT | GCGTATGTTG | CTGTGAATGT | TAAAAATTCA | TCTTGGACTA | 1980 |
|----|----------------|--------------|------------|------------|--------------|------------|------|
| | CCATGAAATC | TAATTTTTCA | AACGCAGCTT | GTACAAAATT | AATATTTGAA | TCCACAATAC | 2040 |
| 5 | CCGTATCTTC | ACCATATAAG | TACAATGAGT | GTACTTCTCC | GTCATGTATA | CCTTCTACCA | 2100 |
| | TTTCATGATT | ATCTTTACCA | GCTTTTGGAT | TCAATTTAAC | GCCATATTCT | TTTTCAAATT | 2160 |
| 10 | TAGCGCGAAT | ATCATCCGCT | TCAATACTTT | GATAACCAGT | AATCTTATCA | GGCATACTTC | 2220 |
| 10 | CCATATCACT | ACATCCTTGA | ACATTATTAT | GTCCACGTAA | TGGATACGCA | CCAGTACCAG | 2280 |
| ٠ | GACGACGATA | ATTACCTGTT | ACTAATAATA | AGTTTGAAAT | CGCTGTACTT | GAGTCACTAC | 2340 |
| 15 | CAATGTCTTG | TTGTGTAATA | CCCATTGCCC | AACAAATTAC | AACAGATTCA | GCTTTAGCAC | 2400 |
| | ATTCTTCAGC | AAATTTAATC | AATTCTGATT | CAGGAATACC | TGTTGCTTCT | TCAGCAAAAG | 2460 |
| | CCATTGTAAA | TGTTTCTAAT | GATTTGTAAT | ATTCATCAAA | ATCATCTACC | CACTCATCAA | 2520 |
| 20 | TAAATGCTTT | ATCGTGTAAA | TCATGATCAA | TAATATACTT | AGTCACTGCA | CTTAACCACG | 2580 |
| | CTAAATCCGT | ACCTGGTTTA | GGTTGATAAA | AACGATCCGC | ACGTTCTGCC | ATTTCATGTT | 2640 |
| | TTCTAATATC | AAATACATGT | ATTTTTTGAC | CAAATAATTT | TTGTGCACGT | TTCATGCGTG | 2700 |
| 25 | ATGCGATAAC | TGGATGAGCT | TCGGCTGTAT | TAGTACCTAT | CAATACAGAC | ATTGCCGCTT | 2760 |
| | TTTCTAAATC | TTCAATACTA | CCTGAGTCAC | CGCCGTGTCC | AACCGTTCTA | AATAAGCCTT | 2820 |
| 20 | TTGTTGCAGG | TGCTTGGCAA | TATCTTGAAC | AGTTATCAAC | GTTATTTGTG | CCAATAACTT | 2880 |
| 30 | GTCTTGCTAA | TTTTTGCATT | AAATACGATT | CTTCATTCGT | CGCTTTAGAA | GAAGAAATGA | 2940 |
| | ATGATAGTGC | ATCTGGGCCA | TGCTTTTCTT | TAATAGCTGT | AAAATTATCT | GCAATGACGT | 3000 |
| 35 | TTAAAGCTTC | ATCCCATTCT | ACTTCATGGA | ACTCACCATT | TTTCCTTACT | AGTGGTTTAG | 3060 |
| | TTAATCGTTG | ATCTGAATTA | ATATGTCCCC | ATGAAAACTT | ACCTTTAACA | CAAGTCGCAA | 3120 |
| | TTTTATTTGC | TGGAGAATCA | TGTGATGGTT | GTACTTTTAA | AATTTCTCTA | TCTTTAGTCC | 3180 |
| 40 | AAACTTCAAA | TGAACAACCC | ACACCACAAT | AAGTACACAC | TGTTTTAGTT | TTCTTAATAC | 3240 |
| | GCTCTTTACG | CATTTCTGCT | TCTGAATCTG | AGATTGCAAA | TAGTGGACCA | TAACCAGGTT | 3300 |
| | CTGCTTTTTT | AGTTAAATCA | ATCATTGCTG | CTAATGAACC | AGGTTCCGTA | TCAGTCATAT | 3360 |
| 45 | AACCCGCATT | ACCTTCCATA | TTCACTTCCA | TCATGGCATT | ACATGGACAT | ACCGTCGCAC | 3420 |
| | ATTGACCACA | AGATACACAT | GAAGACTCAT | TAATCGGTAC | ATCATTATCC | CAAATAACAC | 3480 |
| 50 | GTGGATGTTC | ACGATCCCAA | TCAATTCTAA | TAGTTTCATT | CACTTCGATA | TCTTGACATG | 3540 |
| 50 | CTTCTACACA | ACGCCCACAT | AAGATACATT | GATTTGGATC | ATAACGATAA | AATGGGCCGT | 3600 |
| | N N COMMISSION | CONTROCCORTO | | CARA COMMO | 377777773377 | CCCCAMCCAM | 3660 |

| | TATGCTTTTC | TAAAATTCGA | TCAAGCGCTT | CTTTTTGAGC | ATCTTTCACA | TCATTGTTCA | 3780 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | CAGTATTTAC | AGTCATTGGA | CGATCAATCA | CCGTACTACA | TGAACGTTCA | ATTTTACCGT | 3840 |
| 5 | CAATCTCAAC | AGTACATGTA | TCACATGTTT | GAATTGGTCC | CATCGACTCG | TTATAACAAA | 3900 |
| | TTGAAGGTAC | AAAAGTATCT | TGTGATTTAA | TAAATTCAAG | TAAATTCGTA | CCTGGTTCTA | 3960 |
| 10 | CAAGATAATC | TTTTCCATCA | AGTGTAACCA | CCAAATGTTC | TTGCATATTA | CTCACCCCGT | 4020 |
| 10 | CTATATATAT | TTTCCGTAAA | TGACTTTTAA | TAAATTGCTC | ATATCCACCT | AAAATAACGA | 4080 |
| | TGCCCCACAC | ATCTTTCAGA | TAGAATTAAT | TTAATTGTAT | TACTTTATGT | ACTAGTTGTT | 4140 |
| 15 | AAGTAAAATT | TIGTATITIG | CCTTTTTACA | ATCATTTTTA | TTTGAAATAT | TTTGCGCGAA | 4200 |
| | ATTAAATCAT | CTTTTTGTTT | aattgaaaat | AATTATCATT | ATTAGTTTTC | CAATTATCTG | 4260 |
| | TTTCACGCTT | TTTGCCATAT | CTTTCACAAC | CTTATTAATG | ACAATATTTA | ATAATCACCT | 4320 |
| 20 | CACCTAAAAA | TCGTTATACT | TAAATTTA | ACCCTTTTTC | TGAAAATTAA | TAACCCAAGT | 4380 |
| | TTGATAAATA | TCTACTATCA | TTTAGAAGGT | AATATTTATC | TTTAAATTAA | ATTTGTAATG | 4440 |
| | GATTAATTTA | TAAAAATCAA | ATCAGGCATT | ATAAAATA | GCCCATAAAT | ACAAAGTGTT | 4500 |
| ?5 | ATCACCTTCT | ATTTACGGGC | TATTAGTTCT | ATTCGTTATT | CTATTTACAG | ATCATTCTAT | 4560 |
| | CTAATTAATT | TGTGTACAAT | TTTGATAACT | TATTTTCCCT | TAGTTTACTA | CTCTAGATTA | 4620 |
| 20 | TCTTTTAATA | ACTTAGTACT | TTCAGCTTTT | GACTGCTCAC | TAGGAATGAA | GTAGTACAAT | 4680 |
| 30 | CCGTCACTTT | GAATGCCGCC | TTGACCACTC | AATTGATGTT | TATTAATCGT | GTCATTAGCA | 4740 |
| | TCTTTATAAT | TGCTTCTAAT | CGTATTCAAA | TCACCTAATG | TTAAATCTGT | TTTAACATTA | 4800 |
| 35 | TTTTGAATTT | CATTCATTAG | ACTATTAAAA | TGTGTAATCG | ATGATGGGCT | TGCAATCTTA | 4860 |
| | TTGGCCATCG | CTTCAAGCAC | AATTTGCTGA | CGTTGTTGTC | GACCAAAGTC | ACCACCAGCA | 4920 |
| | cciiciicii | TACGACTTCT | AATAAACTTC | AATGCTTGAT | CACCATTTAC | ATGTGTCTGC | 4980 |
| 10 | TGTCCTTTTG | TAAAACGAAC | ACCATCAACA | GTGAATGTAT | CATTACTTAC | TACATCAACA | 5040 |
| | CCGCCGATGC | TATCTATCAT | ATTATGCAAA | CCATCCATAT | CGATTGTCGC | ATAATGATCA | 5100 |
| | ATTGGCACAT | TCATTAATTT | TTCAAGTGAT | TTAACAGCCA | TATTTGGTCC | ACCATATGCA | 5160 |
| 15 | TAGGCATGTG | CAATTTTTTC | AGTAGTACCA | CGGCCAACAA | TTTCCGCTCT | TGTATCACGC | 5220 |
| | GGTATACTTA | CTATTTCAGT | TTTCTTCGTT | TTAGGGTTGA | TAGATAAAAT | CATAATACTA | 5280 |
| | tCACTACGCT | CTCCGCCACC | CTTTTTCTTA | CGATCAGCAT | CTGAATCGAC | ACCAAATAAA | 5340 |
| 5 0 | GCGATTGTGA | ATGGATCACC | ATCGTTTAAA | CTCACTTTTT | TATCTCTTAA | TTCTGAATGA | 5400 |
| | TTGCGATCTA | ACGGATTGTG | TATCTTATTA | CCAGTAATAA | AAATTTTAGC | AGCTACATAC | 5460 |

| | GGTAGGCTCA | TTTTACTTTT | AGACGAACGT | TTCAATCCCA | CCACTCCTTT | ACTATTCCTT | 5580 |
|----|------------|------------|------------|------------|------------|------------|------|
| | ACATACTTTG | TCTGTTTTCT | CTATTTATTA | TATAGTAAAA | TAATTTTTTT | ACTATACTTC | 5640 |
| 5 | TGTAGACGTA | TAACTATTTT | TTATCATTTT | TTATCTCTAG | AGAATATCTA | TCTGTATTTT | 5700 |
| | TGATAACCAC | CATTTGCATT | TAAAATTTTA | AGTACCGTTT | CATGACATGC | TTTATTACTT | 5760 |
| | ATAATAAAAG | GTGCACCCTT | TAAATGATCA | ATTGCCTTAC | CATCTAAAGT | CGTCATTTTT | 5820 |
| 10 | AGATTCAATA | GTTCTGCAAA | TAAAAACTGT | GCAGCAATGT | CCCAAGGTTT | AGGATTTGTA | 5880 |
| | TTAATATGTG | CCCCAAATTG | ACCTTTTGCC | ACTCGCATAG | AATCTAATCC | GCAAGCACCA | 5940 |
| 15 | ACTAAACGAT | AACTAAATGA | GGCGTCAAAT | AAATCTTGCA | CCGTATCTAG | ATTCATCACT | 6000 |
| | TGTGCATTAA | ACGATATAAT | AGCGTCTTCC | AATTTTAACG | ATGGTGGTTC | TTCCATCTTA | 6060 |
| | ATTCCATTAC | AAAAAGCACC | TTCTCCTCGT | ATTGCTTTAT | AAAGCTTTTT | ATGCGGATAA | 6120 |
| 20 | TCATATACGT | ACGATAACAT | TGGTTTACCT | TCATAAAAAT | ACGCCAATAT | AATACAATAA | 6180 |
| | TCTTCTTGCT | GTTTTACTAA | ATTGGCAGTT | CCATCAATGG | GATCCATAAT | CCATAAATGA | 6240 |
| | TTAATTTCAT | TCGTAATCAT | TTCATTACTT | TTTTCTTCCG | CTAATAGTTG | GTGTTCCGGA | 6300 |
| 25 | AAATGTGTTG | CTAAAAATTG | TTGGAATTGT | TGTTGAATCT | GTTTATCTAC | ATTTGTAACT | 6360 |
| | AAATCAAATC | GATGACGCTT | AGTTTCTGTA | GTCATTTCCA | TAATTAATTG | CGGAATAACA | 6420 |
| | TTGTCTATTT | GTTTCAACCA | CGAACATATT | AACTTATCTA | TTTGCTGTAA | TGTTTTATCT | 6480 |
| 30 | GTCATTTCGT | CCACCACTTC | TCATATCATT | ATCATTTAT | TATTACCCTA | TATTAAAAGA | 6540 |
| | ATCAACAATA | CAACTGAAGA | CTTCTTCATT | TTATGCATAA | AAAAATCGGC | TAGTCACGTG | 6600 |
| 35 | CTAGCCGACA | AATAGAAAGG | AAAGTAAGTA | ATAAATATTG | AAGATGTTGT | GATGTAACTT | 6660 |
| 35 | GAACGATTAA | AAGCTATCTG | TTATATAGCT | CTACCCCTTT | GTTTAATCGC | TCCCCCTGTT | 6720 |
| | ACAAGTAATA | TCATAGCACA | ATCTTTTTTA | AAATGTAAGC | GTTTTCCACA | AAATTTTTAC | 6780 |
| 40 | GATTTTTTA | AAAAGATATT | GAAAATGTCC | TCATTGTCAC | TCTTATGTTA | TACTTTGTGT | 6840 |
| | AATATATCAT | CTTTTAGGAG | GTGGCTGTCA | TGAATAAAGC | TGAAAGGCAA | AATTTAATAA | 6900 |
| | TTACTGCAAT | TCAACAAAAT | AAAAAATGA | CCGCTTTAGA | ATTAGCTAAA | TATTGCAACG | 6960 |
| 45 | TATCCAAACG | CACAATTTTA | AGAGATATTG | ATGATTTAGA | AAATCAAGGT | GTTAAAATTT | 7020 |
| | ATGCGCATTA | TGGGAAAAAT | GGTGGTTACC | AAATACAACA | AGCACAATCT | AAAATTGCAT | 7080 |
| | TAAACTTATC | TGAAACACAA | TTATCAGCCT | TATTTTTAGT | GCTTAATGAA | AGTCAGTCGT | 7140 |
| 50 | ACTCGACATT | ACCATATAAA | AGCGAAATCA | ACGCAATTAT | AAAACAATGT | TTAAGTCTTC | 7200 |
| | CACAAACACG | CTTAAGAAAA | TTGCTTAAAC | GCATGGACTT | TTATATTAAA | TTTGATGACA | 7260 |

| | ATGTGATGTT | AGTAGATCAT | AGGGTTGATG | ATAATATTAA | AGCTGAAAAC | GTTATATTTA | 7380 |
|----|------------|------------|------------|------------|------------|------------|------|
| | TTGGCCTTTT | GTGTAAACAT | GGACATTGGC | ATGCAGTCAT | TTATGACATT | GCTCAAGACA | 7440 |
| 5 | AAACTGCCGA | ACTCGAAATT | GAAAATATTA | TAGATATTTC | GTATTCATTC | GGTAAGACGA | 7500 |
| | TTCAAACCAG | AGACATATCC | ATTGATAACT | ATCATCAATT | TTTAAACCCC | ATCGATTCCT | 7560 |
| | AAAAAACAGC | AGTAAGATGA | TTTTCAATTA | GAAAATATCT | TGCTGCTGTT | CTCTATTTAT | 7620 |
| 10 | ACAATACTTC | GTATTGAATG | GnTTCGCTTT | CCTAGGGTGC | CGTCTCAGCC | TTGGTCTTCG | 7680 |
| | ACTGGCACTG | CTCCCTCAGG | AGTCTCGCCA | TTAATACTAC | GTATTAACAT | GTAATTTTAC | 7740 |
| 15 | TTTGAAATAC | TTAAAAAAAT | AAAACACTTT | GCCCAACTTA | CACTACCAAT | AGAAACTGCT | 7800 |
| | GTTAGAATTC | CTCAAAATGA | TATTTCGCGA | TATGTTAATG | AAATTGTTAA | AAAGATAGCT | 7860 |
| | GATAGCGAAT | TCGATGAATT | CAGACATCAT | CGTGGCGCAA | CATCCTATCA | TCTAAAAATG | 7920 |
| ?0 | ATGTTAAAAA | TCACCTCATA | TTCATATACT | CAATCTGAAT | TTTCTGGCCG | TAGAATAGAA | 7980 |
| | AAATTACTTC | ATAACAGTAT | TCGAATGATG | TGGTTAGCTC | AAGATCAAAC | ACCTTCTTAT | 8040 |
| | AAAACTATTA | ATCTTTTTAG | AGTGAATCCT | AATACTGATG | CGCTAATTGA | ATCTTTATTT | 8100 |
| ?5 | ATTCAGTTTC | ATAATAAAAT | GCATATCAAA | AAAGCTGATT | TCTATCAAAT | AATTAATAGA | 8160 |
| | AATCAGCTTT | TTTCaTTGCC | TAAAAACTTA | ATGTCCCGAC | CTCTTTATCT | ACGCATAAAT | 8220 |
| | ACTTATTACT | GATATAACGA | AAGAAACAAA | ATTATTTGCT | ATATGTAATG | CAATTGTTGA | 8280 |
| 30 | ACCTAGGTTT | CTTCCAGATT | TTAAATAAGT | GAAAACTAAT | ATGATGGATA | GTATGAGATA | 8340 |
| | TGGACCAAAC | TCAAACGGCG | ACTTTGCATC | AGTCACATGA | ATAAATGCAA | ATAAGAACAC | 8400 |
| 35 | CGAAACAATA | CTCATAGCTA | TAAAATTAAA | CTTCTTACCT | AATTCTCCAA | TTAAAATATG | 8460 |
| | TCTAAATACG | ATTTCTTCAA | CTATTGGACC | TACAATCACA | ATTAATAAGA | ATGCTACAGG | 8520 |
| | TAAAAATGCA | GGCACTTCAA | ACATTTTATT | TAGCTCAAGT | TCATTGGCTG | TTtCACTATA | 8580 |
| 10 | TTGĊAAATGT | TTAGGTAGAA | ACTGTGTCAT | ATATTCATAT | GTATAAATTA | AGATGAGAGC | 8640 |
| | AATAATATAC | GTTATTGACA | ATCTAAGCCA | ATATTTTTTG | ATATACGCAA | AACCAGCTCG | 8700 |
| | AAGCCTTGAT | GGCATCACTT | TTAAATGAAA | TAAATAAAT | GCGCCAATCC | CAATCGTATA | 8760 |
| 15 | TGCTAAAGCT | TGTGTGATAG | TCGCTACAAA | TATCAGATTA | CTATCGATTT | CATAATAACC | 8820 |
| | AAACAAAATT | GGTCCTATGT | AAGCTGCAAT | TGTGAGTGCA | TAAAATATAA | CACCTATAAT | 888 |
| | TGGAATTATA | AGCAAATCTC | TCCATGCTAT | ATCTTTAAAC | GTGTATTTCT | TTTTTTCATT | 8940 |
| 50 | TTCCaCTGTT | ATATCCTTTC | CTGTTTAATA | ATTGATTTTT | GGAGGTACTT | CTACATGATA | 9000 |
| | AACGAAACTA | AGTATATGAG | ACAACAAATT | ACTAATTTGA | TTCABATCAT | TGATACGATT | 9060 |

| | ATAGTTACTA ATGAATTGAA TAAGTTCAAA GGCTTTGAAA CATCATATAT AATAAACGAA | 9180 |
|-----------|---|-------|
| | AATCAAGTTT CCTATTATGA AATTATAACA CTACTTAATA AACGTCCCCT CGACAAGTCG | 9240 |
| 5 | ACTATGGTAA CAAAATTCAA TATCTTAATT TTTATCATAC AGAACTATCT AACGCATTAT | 9300 |
| | TTGCAATTAA ATTTGCCCAT TAACCTATTT TTCATAAAAT GTCATTTAAA CAAGTTATTT | 9360 |
| 10 | ATTAAAATTC ACTTTATTAC ATAAATTATA CAATTAYAAA GTTTCTTCAA ATTGTAAAGA | 9420 |
| ,, | TGCATTAATC GAGTTATAAT CATAATGATT AAGATGGT | 9458 |
| | (2) INFORMATION FOR SEQ ID NO: 115: | |
| 15 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 910 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 20 | | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 115: | |
| | ANGCGTATCA TGTCACGCAT TTTAACTACT TCTTTACCAC AAGATTATAC AGTCACATTA | 60 |
| 25 | GTTGATCGTA TGCCATTTCA TGGATTGAAA CCAGAATTTT ATGCTTTAGC TGCGGGCACG | 120 |
| | AAATCAGATA AAGATGTTCG TATGAAATTC CCTAATCATC CACAAGTGAA TACAGTTTAT | 180 |
| 30 | GGTGAAATTA ACGACATAGA TTTAGATGCT CAAATTGTCT CAGTCGGTAA TTCTAAAATT | 240 |
| | GATTATGATG AGCTAATCAT TGGTTTAGGA TGTGAAGATA AATATCATAA CGTTCCAGGA | 300 |
| , | GCCGAAGAAT ATACACATAG TATTCAAACA CTCTCAAAGG CTCGGGATAC TTTCCATAGT | 360 |
| 35 | ATTAGTGAAC TACCAGAAGG TGCTAAAGTC GGTATCGTTG GTGCTGGATT AAGCGGCATA | 420 |
| ٠ | GAACTTGCCA GCGAATTAAG AGAAAGTAGA TCAGACTTGG AAATATATCT TTATGACCGT | 480 |
| | GGGCEGCGAA TTTTAAGAAA TTTTCCAGAA AAATTAAGTA AGTATGTTGC GAAATGGTTC | 540 |
| 40 | GCCAAAAATA ATGTTACCGT TGTTCCAAAT TCAAATATTA ATAAAGTTGA ACCTGGTAAA | 600 |
| | ATATATAACT GTGATGAACC TAAAGATATT GATTTAGTTG TATGGACAGC AGGAATTCAA | 660 |
| 45 | CCTGTTGAAG TTGTTCGTAA CTTGCCGATT GATATAAATA GTAATGGACG CGTGATAGTT | 720 |
| | AACCAGTATC ATCAAGTACC AACATATCGT AACGTCTATG TAGTTGGTGA TTGTGCTGAT | , 780 |
| | TTACCACATG CGCCAAGTGC TCAGTTAGCC GAAGTTCAAG GTGATCAAAT TGCCGATGTG | 840 |
| 50 | CTTAAAAAGC AATGGCTAAA TGAACCATTA CCTGACAAAA TGCCGGAACT AAAGGTACAA | 900 |
| | GGTATCGTTG | 910 |
| | (2) INFORMATION FOR SEQ ID NO: 116: | |

(A) LENGTH: 10182 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 116:

| 10 | TTTTTGATTC | AAAGTGGTGA | TTTAACAAGC | ATTTTAAATA | GCAATGATTT | GAAAGTCACA | 60 |
|----|------------|------------|------------|------------|------------|------------|------|
| | CATGATCCTA | CCACTGATTA | TTATAATTTA | TCTGGTAAGT | TGTCGAACGA | TAATCCAAAC | 120 |
| | GTTAAACAAT | TAAAACGTAG | ATATAATATT | CCTAAAAACG | CATCAACAAA | GGTGGAATTA | 180 |
| 15 | AAGGGAATGA | GTGATTTAAA | AGGCAATAAT | CATCAAGATC | AGAAACTTTA | TTTTTATTTT | 240 |
| | TCAAGTCCTG | GAAAAGACCA | AATCATTTAT | AAAGAAAGCC | TTACTTATAA | TAAAATAAGT | 300 |
| 20 | GAACATTAAT | ACTTATGCTG | TAATTATAGA | AACATCCAAA | TCATCTATTA | nAATCCTATA | 360 |
| 20 | TTATAAAAnC | ACCTCACATA | ACTCGTTCAA | CTGTACCAAA | CCACATTACA | TTAGATTTTA | 420 |
| | GGCTAACTAT | TGTGATGTAC | ATCAAAAACG | AATTTGTGAG | GCGTTGTATA | TTTTACAAAG | 480 |
| 25 | GTGACTAGCG | TTTCGTATAG | CATTTCCAAC | ATTACTACAC | TCAAGCGTCA | CGCTAAAGTT | 540 |
| | CGAAATCGAA | TCCTTTCATT | CAACAAAAGC | TCATATCCAC | TACAAACTTC | ATATCAAGCG | 600 |
| | TATAAACTAT | CTTGTGATAC | TATCTCGATC | ATATCTATAG | TATGCATTTG | TGTTCCGTTT | 660 |
| 30 | CACTGAAGTA | TATGTATCAT | CAGTTAAGTA | TAAACCGTCA | TCCTTCAATG | TTACTTGATA | 720 |
| | AGCATATTTC | CGTGCTAACC | AGGCAATATC | TATATAATTT | TCTCCTGCGT | TTTCATAACT | 780 |
| | TCTTAAATCT | TCAATATGTG | CACTAACTTC | AGGGAAAATG | ATTCTAACAA | CACTTTCATC | 840 |
| 35 | AACCCAATAT | TTGTCATGCA | TCCATCGCAC | TTGATCTGCC | AATAAAGGTA | ACTGCACATC | 900 |
| | ATTGAAATAT | AGACGAAAGC | CGTCACTATC | ATACATTTGC | CGATATGGTA | ATGGCTGTTT | 960 |
| 40 | TCTAĀTCACT | AACACCTCGC | CACCCATTAC | GGTGCCTTCT | CTAGTATCAT | CACTTCCACC | 1020 |
| 40 | CGAAGCTTCA | TACGTTGTTG | GGTCAACCTG | TAGTCCATGT | ACATCTCCAA | TATAAGCATC | 1080 |
| | TGGTTTATGT | TCCATTGCAT | GTCCATGTGC | AATCAATGCT | AATATTGTAG | ATTGTGAAAA | 1140 |
| 45 | TTGAGGCTCC | CATTCAATGC | GATTAGGATG | GCTACTATAA | ATTCTAGGTT | CATCTATAGC | 1200 |
| | CTGCTGAATA | TCCATGCCAA | ACACTAATAC | ATTGATTAAT | GTTTGCGCAA | CACTAGCAAT | 1260 |
| | GATACTTATG | GCACCAGGTG | CACCTACTGT | TAATATTGGC | TTCCCGTGAT | ACATCACAAT | 1320 |
| 50 | CGTTGGAGCC | ATGTTACTTA | GTGGTCGTTT | ATATGGTGCA | ATTTCGTTAA | TACCACCATC | 1380 |
| | TACTACATCA | AAGCCATCCA | TTGTCGTATT | CAATAACACA | CCGTAGCCTG | GAATCGTGAT | 1440 |
| | ACCTGAACCA | TAAATCATAC | CAATTGATGT | CGTAAATGAA | GCAATATTAC | CTTCCTTATC | 1500 |

55

| | ATCAGACACA | ACACCATGCT | CTATATCAAT | ATTTGCTTTA | TTGCTATCAA | TGAGCGTACT | 1620 |
|----|------------|------------|------------|------------|------------|------------|------|
| | GCGTGCTTTT | AAATAATCAT | CATCAATTAA | TGACTGTACA | GGCACCTCAT | GAAAATTATC | 1680 |
| 5 | ATCCGCCAAG | TATTGCGCAC | GATCACTATA | TGCTAAATGC | ATCGCTTGTA | TCAAATGATG | 1740 |
| | CAAGTAATCA | ACAGATCTTG | GACCCATAGA | TGGTAAATCG | ACATGTTCTA | ATAACTTCAA | 1800 |
| 10 | TATTTGAATT | ACCGTGATAC | CGCCAGAACT | AGATGGTCCC | ATTGAATAAA | TGTCATAGTC | 1860 |
| | TTTAAATGTT | GCACTGATTG | GCGCTTTAAT | CTGAATGTCA | TATTTGGCTA | GATCCTCTAA | 1920 |
| | AGTGATTGTC | CCACCACATG | CTTTGACAAC | ATTGACTAAT | TGTTTCGCAA | TGTCACCTTT | 1980 |
| 15 | ATAAAATGCA | TTAAACCCTT | GTTCTCTTAA | TATTTGAAAT | GTCTTACCTA | ATTCGGGTTG | 2040 |
| | TACAATCCAA | TCACCTTCAC | GCCAATATTG | ATTTTCATGC | GTAAATACTT | GTGCCGTTTC | 2100 |
| | ATGATACTTT | GTCAATCGTG | CGTGTTGCTG | GCGCGAATAT | TTTTCAGTAG | CCCAATTGGC | 2160 |
| 20 | TGCATGACCT | TCAATGGCTA | GTTCAATTGC | AGGATTAATT | AAATCTTCCA | ATGACAATTT | 2220 |
| • | AGCATAACGC | TTGTGAATAT | AATCAAACAG | CTTTGGAATT | GCTGGCACAG | CGACAGTTTT | 2280 |
| | ACCATGTGTA | GTCATATCAA | AAAATGATTT | ATATTCGCCT | GAATCATCTA | GATAAAATTG | 2340 |
| 25 | TTTGTCTACA | TGTTCAGGTG | CTGTCTCACG | TGCATCAAAC | GCAGTTATAC | TGCCAGTACT | 2400 |
| | TTGCTCATAA | TATAGCAAAT | ACCCGCCACC | ACCAATACCT | GATGCAAATG | GTTCTACCAC | 2460 |
| 30 | ATTCAATGCC | AGTTGAATTG | CAATCACTGC | ATCCATGGCG | TTGCCACCTT | GATCTAATAC | 2520 |
| | ATCCTTACCA | ATTTTAGCCG | CAAGAGGATG | TGATACGGAA | ATTAACCCTT | CTTTAGATGT | 2580 |
| | TTTTGTCTGT | TTGTCATTTA | AGTTAATGAC | CATACTATAT | CCTCCTACTT | TCTGTTAAAT | 2640 |
| 35 | ATTTAAAACA | TTATTGATTA | ATGGCTTTTT | CTACTTTTTC | TAAATCTTGA | CGTTGCTCGT | 2700 |
| | TACCAGTATC | GACAAGTGGT | GTAATCGGTG | ATGCAATTTT | AAATTTATCG | CCACGATAAA | 2760 |
| | ACTTAATAAA | TTGATCCTGA | TCTATCGCAT | TAACTACTGC | TTGTCTCAAG | TTTGGATGCG | 2820 |
| 40 | TCTTAAATAT | ACCTTTTTTA | ATATTTAGCA | TTAAAAAGAC | TGACTTGCGT | CCATTTTTGC | 2880 |
| | GAATAATGCT | TAAATTTTTA | TCCGACTTAA | TTAAATCAAA | ATGTTTTTGA | TTCACATCTG | 2940 |
| | CCAACATATC | aattgaatga | TTTCTAAGTT | CTGACAATGC | ATTATTCGGG | TCACCATTAA | 3000 |
| 45 | ACTTCAATGT | AATATTTTTA | ATTTTAGCTG | GTCCATAACT | ACCTTTTTCT | GTTTCGTTGA | 3060 |
| | ATCCTGGATT | ACGTTGAAAC | GTTGCTTGAT | ATGCATTTTT | CTGTGTCATA | ATGTATGCGC | 3120 |
| 50 | CACTTGCATA | CAGCGCATTT | TTCCCATCTG | AATTTGCAGG | AATTGTACTG | CTATCCCCAT | 3180 |
| | ATCCTTTTGG | ATATTCTTGA | TTTACTTGAT | TAACAAATTT | TTTAGATAAA | ATGCCTGCCG | 3240 |
| | AAGAGTGTGT | TAAGTAATTT | ACCTCTCGAG | GCATCGATTG | ATCTGTCGTA | ATTTTAACAA | 3300 |

| | TATAAGCTTT | AATCAACTTA | TCATAGATTG | ATTTATCGTC | CITGTCTTTC | TCTTTACGCA | 3420 |
|----|------------|------------|------------|------------|------------|------------|------|
| | ACTGATCGAT | GTCCTCATCT | TTTAATATCT | TGATGTCATT | TATATGTTTG | TGCATATTGT | 3480 |
| 5 | AAGTATTATT | GTTAGGCACA | GACTTTTTAT | CACGTGCTCT | ATCTAAAGAA | AACTTAACAT | 3540 |
| | CTTCAGCCGA | TACACGCTCT | CCAGTATTAC | GTGCTTGTCC | ATTGACCACT | TTCGCAAAAT | 3600 |
| | AATCATCATC | TCTTAACAAG | AAATAAAATG | CTTTATTGTC | CTTATTCACA | GCATAATCAT | 3660 |
| 10 | GACTTAACGA | ACCTTTCGTT | GTTAAATGAT | CATTTTCATC | TAATAATAAT | AACCTTGTGT | 3720 |
| | ACATATTCAT | attaattgaa | TATACTGACG | GCGCAATTGA | ACGTATTGGA | TCCAATGTAG | 3780 |
| 15 | GAATTTCACC | ATCTTGTTGT | GTCATCACAA | GTGGCCGCGT | ATCTCGTTCT | CTACTATTGT | 3840 |
| | TGTAATCAAA | TTGTTGCCAT | ATTAATGCAC | GTGAATTTGG | CAATCCAACA | CTATTTTTAT | 3900 |
| • | CTAACACTTT | ATTGTCATAT | ACTAAATTCT | TTTTTGATCC | ATATAAAGGC | GCCATATACC | 3960 |
| 20 | CTTTATCAAA | TACAACTTCA | TCTTCAATTT | GCTTATATGT | TTGTTTAACA | TCTGCTTCAT | 4020 |
| | TTTGAGTAGA | AGCTTTATTT | AACAACTGGT | CTACATGTTT | ATCTTTCAAT | AAACTATTTG | 4080 |
| | ATCCTGTAGA | ACTAAATAAT | GCCGTCATAG | CATAGTTCGG | GTCACCAAAC | ACTGTCATCC | 4140 |
| 25 | AGTCATCAAT | TTGGATATCA | TAATTGCCGG | CTTGACGTTG | TGTACGATAG | CTACCATAAT | 4200 |
| | CTGGTTGGAT | ATTCATCTTC | ACGTTAAATC | CTGCATTTTC | CAATTGATCT | TTAACGATAT | 4260 |
| 30 | TCATATCATT | TTCATAACTT | GCTTGTCCTA | GGAAATGTAT | TGTTGGTCGC | TCGCCTTTCA | 4320 |
| 30 | CTTCAACTTT | CGATGACTTT | TGAGCCACTT | CTGATTTCGT | AGGGACACCA | CAACCACTTA | 4380 |
| | ATACCAACGC | таааастата | ATTGCGATAC | TAATGATTTT | CTTCACATCT | ATCCCTACCT | 4440 |
| 35 | TTTTAATGAA | TTCTTGGATC | TAGTGCATCA | CGCACTGCAT | CACCTATAAA | ATTAAATGCT | 4500 |
| | AAAACGACGA | ACATAATACA | AACACCAGGT | ACAATAGCTA | AATTACTGTG | CGTTTCCAAG | 4560 |
| | TAGTTACTAC | CGGTACGTAA | AATGTTGCCC | CATTCAGCTA | CATCAGGTGC | AACACCAAGT | 4620 |
| 40 | CCTAGGAAAC | TTAAACTACT | TGTTGTTAAT | ACAACCACAC | CTATATITAA | TGAAAAACGT | 4680 |
| | ACAATCATAG | GCGCAATCGC | ATTCGGTAAA | ATATAACGCC | ATATGATATT | CCAAGTGTTT | 4740 |
| | TCACCAGTGA | TACGTGCTGC | ATCTACATAT | TCCATGCGTT | TAATTTCTAA | AACACTGGCA | 4800 |
| 45 | CGCATTGTCC | GTGCAAATGA | TGGTATATTA | CCGATACTTA | AAGCAATAAT | TAAATTTGGA | 4860 |
| | ATACTTGCTC | CAAATGATGC | AATAATTGCC | ACCGCTAACA | ATAATGATGG | AATTGCAAAC | 4920 |
| 50 | ACTACATCTA | AAATTCGCAT | TATTAAATTA | TCAATATGAT | TAAAATAACC | TGCGATAGTG | 4980 |
| | CCTAGTAACA | CACCAAAAAT | AACTGCAATA | ACTACTGAAA | TAATTGAAAT | TGAAAATGTC | 5040 |
| | AGCTTCGTTC | CTACAACTAC | GCGTGTAAAT | AAGTCTCTAC | CGAAATCATC | AGTACCAAAC | 5100 |

Ť.:

| | GTATCAAATG | TAAATTGTGA | CACAATTGAT | AATGTCAGCA | TGTAGACTAA | AATAAGTAAC | 5220 |
|----|------------|------------|------------|-------------|------------|------------|------|
| | CCGATAATCG | CAATACGATG | TCTAGTAGTT | TTTCGTATAA | ACGATTCCCA | CCCGTTATAA | 5280 |
| 5 | CTATGTATTT | GCGATGTACG | TTGGTAACGT | СТААТАСТТА | CAAACATTAA | TAATGTAAAT | 5340 |
| | ACGTTGCCTG | TTAATGTCAT | CAACAATAAC | AACACTTCGA | CGATACGTCG | CCATAGGTCA | 5400 |
| | TGATGCTTCC | ATGTTTGTTC | CGTTGTTAAA | ATAATAATTA | AAATGATGGT | TAAAACGATT | 5460 |
| 10 | AGCAATGTTT | CAGCAATATA | GAACGTATCG | GCCACATAAC | CTTTAAAAAG | ATTTAATGCA | 5520 |
| | CTCGTTAATA | TAACTAAAAT | ATAAGTTGCT | ATGGCGTAAC | TTGCGAATAA | TTTTAAGGAA | 5580 |
| 15 | GCTATCTTTG | AATTAAGTTG | TGCCATATGC | CTCACTTCCT | TTCGTTGATT | TCACTACGTA | 5640 |
| | ATTTTGGATC | GATTAAAGCA | TAAAATATAT | CAATAATTAA | GTTTGCTAAA | GATATTACAA | 5700 |
| | TTGATATATA | TACGACCCCA | CCCATGACTG | CTGGAATATC | AGGTATTAGT | TGTTTTTGGA | 5760 |
| 20 | CGATATAACG | CCCGATACCA | TTAATGTTAA | ATACTTGTTC | CGTCACTGCT | GAACCGCCTA | 5820 |
| | GTAACTCTGC | CACTAGAAGA | CCAACTAACG | TTACAATTGG | AATAATGGCA | TTTTTCAAAA | 5880 |
| , | TATGTTTAAT | AACAACTTGT | GTCGTCGATA | ATCCTTTTGC | ATAAGCAGTT | AAAACATAAT | 5940 |
| 25 | CGCtGCGCAT | TACTTCAAGT | ACAGAAGACC | TTGTCATACG | CGTGATAGAA | GCAGCAATAC | 6000 |
| | TTGTTCCAAT | GACAAGTACA | GGTAAAATCA | ACGATATTGG | ATGTTCTGGC | ATATAAGATG | 6060 |
| 20 | GTGGCAAAAT | ATCCAATTTC | AATGAGAACG | CTAAAATGAA | TAATAGCCCT | TGCCAGAAAC | 6120 |
| 30 | TTGGAATAGA | TAAACCAATT | AATGCAATTA | TCATTAACGT | GATATCAAGC | CAACTATTTC | 6180 |
| | GCTTCATCGC | ACTGATAATA | CCAATTGGTA | TTGCAATAAT | TAATGCCACC | ATTAGCGCTA | 6240 |
| 35 | ATACTGCGAC | AATTATTGTA | ATTGGAATTC | TTTCGCCAAC | TGCTTTAGTC | ACAACCTCAT | 6300 |
| | TCCCTTTGTA | AGTCGTACCT | AAGTCAAAGG | TAAAAACACC | CTTGATGGTA | TCCCACAATT | 6360 |
| | GAATAAAATA | AGGTTCGTTA | AGATGATGTA | ATACATTGAA, | TTGATGTATC | TGTGCCTTTG | 6420 |
| 40 | TTGCATTTTG | TCCCAGTATG | CTATAAGCCG | CATCAAGCGG | TGAAAAATAC | AGAATGGTAA | 6480 |
| | ACACACTGAC | AATAACACCA | ATGATGACAA | TCACAGCCAT | GACAATTCGT | TCAAAAATAT | 6540 |
| | ATCTAACTAA | TGGCTGTAAA | TAAAAAGTCA | ATAAGATGAA | CATCGGCAAG | GCCAATATCA | 6600 |
| 45 | CTTTGATCAT | GATGAACTTA | TGAAATAATA | CATTTTCAAA | GTATGTTGAA | AAATGTGCTT | 6660 |
| | GTTCAATATT | CTTTGAACTC | GTATTAGAAC | TTTGTGCCTT | GAATATTTTT | AATGCTTCTT | 6720 |
| | TATGTATTTG | TGTGGATGAC | TTTTGCTGCG | ATAAATATTT | ATATTTTTGA | TGTAACGCCT | 6780 |
| 50 | GTTCAATTTC | TGAAATTTCA | GAATTATTAG | CGTAAAAATT | TTTCCTCTTA | GCAGAAAAGA | 6840 |
| | AAAACTTTAT | CACTGCATAT | AAAAATATTG | GCAAGCTTAA | TACCGATAAT | ACAAACTTGT | 6900 |

| | CTTGTAAAAT | AATCTTGAGT | AGATTACTAT | GATATACAAA | AGTATAGAAT | AAATTTACAC | 7020 |
|------------|------------|---------------|------------|------------|------------|--------------|------|
| | ATTTGTGaAT | AGGGAGGCAC | AACATCATGT | CAAATTTATT | AGAAGTCAAC | AGTCTGAATG | 7080 |
| <i>5</i> | TACAATTCAA | TTATGATGAA | ACTACAGTTC | AAGCGGTAAA | AAACGTCTCT | TTCGAATTAC | 7140 |
| | GAAAAAAACA | TATCCTAGGT | ATTGTTGGTG | AATCAGGATC | AGGAAAAAGT | ATTACCGCTA | 7200 |
| 10 | AATCTATTTT | AGGGCTACTA | CCAGATTATC | CAGATCACAC | ATTAACAGGA | GAAATTATTT | 7260 |
| ,,, | TTAATGGGCA | ATCGTTAAAT | AATTTATCAA | CTTCAGCGTT | ACAACAAATT | CGAGGTAAGG | 7320 |
| | ATATTTCAAT | GATTTTTCAA | GATCCACTCT | CTTCGTTGAA | TCCAAGATTA | ACGATTGGCA | 7380 |
| 15 | AACAAATTAC | AGAAGTAATA | TTTCAACATA | AACGTGTATC | TAAATCTGAA | GCAAAGTCGA | 7440 |
| | TGACAATAGA | CATTTTAGAA | AAAGTAGGTA | TAAAACATGC | AACTCGACAA | TTTGATGCTT | 7500 |
| | ATCCACATGA | ACTTTCTGGT | GGTATGCGTC | AACGTGTCAT | GATAGCAATG | GCATTGATTT | 7560 |
| 20 | TAAAGCCACA | AATTTTAATC | GCAGATGAaC | CAACAACGGC | ATTAGATGCC | AGTACACAAA | 7620 |
| | ATCAATTACT | GCAGTTAATG | AAGTCCCTTT | ATGAGTACAC | AGAAACATCT | ATTATTTTA | 7680 |
| | TCACTCACGA | TTTAGGCGCT | GTGTATCAAT | TTTGCGACGA | TGTGATTGTA | ATGAAAGATG | 7740 |
| ?5 | GAAGTGTCGT | TGAAAGTGGC | ACGGTTGAAA | GTATTTTTAA | ATCGCCACAA | CATACCTATA | 7800 |
| | CAAAACGCTT | AATAGATGCG | ATTCCTGATA | TTCATCAAAC | GCGTCCGCCA | AGACCGTTAA | 7860 |
| 3 <i>0</i> | ACAATGATAT | TTTATTAAAA | TTCGATCGCG | TGAGyGgGAT | TACACATCAC | CGAGTGGCAG | 7920 |
| | CCTATACCGA | GCAGTTAATG | ATATTAACTT | GGCTATTAGA | AAAGGCGAAA | CATTAGGCAT | 7980 |
| | TGTCGGTGAA | TCAGGGTCAG | GGAAATCGAC | ATTAGCTAAG | ACGGTCGTCG | GTCTAAAGGA | 8040 |
| 35 | AGTGTCAGAA | GGCTTTATTT | GGTATAACGA | ATTACCATTA | AGTTTATTTA | AAGATGATGA | 8100 |
| | ATTGAAATCT | TTACGACAAG | AGATACAAAT | GATTTTTCAA | GATCCATTCG | CATCTATTAA | 8160 |
| | TCCAAGATTT | AAAGTCATTG | ATGTGATTAA | ACGACCACTA | ATCATTCATG | GGAAAGTCAA | 8220 |
| 10 | AGATAATGAT | GACATTATTA | AAACTGTCGT | ATCGTTGTTA | GAAAAGGTTG | GCCTAGATCA | 8280 |
| | AACTTTCTTA | TATCGCTATC | CACACGAATT | ATCTGGTGGG | CAACGTCAGC | GTGTAAGTAT | 8340 |
| | CGCGAGAGCA | CTTGCTGTTG | AACCTAAAGT | GATTGTTTGC | GACGAGGCAG | TGTCCGCTTT | 8400 |
| 45 | AGACGTTTCA | ATTCAAAAAG | ATATCATCGA | GTTATTAAAA | CAATTACAGT | TAGACTTCGG | 8460 |
| | CATCACTTAT | TTATTCATCA | CACATGACAT | GGGTGTTATC | AATGAAATAT | GTGATCGCGT | 8520 |
| 50 | TGCAGTTATG | AAAAATGGCG | AAATCGTTGA | ACTGAATAAC | ACAGAAGATA | TTATCAAACA | 8580 |
| | TCCGCAGTCA | GACTATGCAA | AGCAACTTAT | TTCAGAAGTA | GCAGTTATTG | СТАААТАААА | 8640 |
| | CTCATCCTT | CTCC A A CTTT | ATCACTCTAT | CCTCTCDAAT | 222000000 | N COMPONENTS | 0700 |

| | TATCAAGTTT | TAGGTGCTTT | GCCATGATTT | AAGAGTCACC | CCCATACTTT | GGGCATTTTA | 8820 |
|-----|------------|------------|------------|------------|------------|------------|-------|
| , i | ACGCCAGAAT | AAATCCCCCG | CCACTATGTG | AAGTGTGGGG | GATTATTTAT | ATTTTATTAG | 8880 |
| 5 | AATATTCAGA | TTTTTGAGTG | TGTCAACTTA | GCTTAGTCAA | TGTATATTTA | ACGTCACTTA | 8940 |
| | CTCTTTTTCT | TTCATAATTA | ACACATTCAA | ATAAACTTTG | ATCAAAAAAC | ACAAAGTTAA | 9000 |
| 10 | AAGTACCATC | TTGTAATATG | CTCTCATACA | TTATCCCGTC | ATATTTAAGG | CTTCGAATAT | 9060 |
| • | AATCAGCTAA | ATATTGAAAT | GGCAAATAAT | CTATTCCTTG | TTCATCGCTT | GGATTTGTTA | 9120 |
| | TTCCTTTATG | AATCTTTTTT | AATGTTTGGT | AATTTACAAA | ATACTTTCTA | AATCCATCAT | 9180 |
| 15 | CGCCAGCTTT | GATTGCATTA | CTAGTTAAAT | TAGTTAAATT | CGCAATTTTC | AATTTCTCTT | 9240 |
| | TTGTCACGTT | TTTTTGTAAC | TTAACCTTAC | CTATATAAAT | AATGTCATTA | TGCTTAGGTT | 9300 |
| | TAACTTCTTC | TATACTGACC | TGTTCTTTTG | TACTAAGGTA | TAATACGCTT | ATCCATTTAG | 9360 |
| 20 | AATTCAATCT | TCCTGCCGTT | GCAAATCCCT | TTGGTGGTGA | CATTAGTTCA | CTTTTCTCTG | 9420 |
| | TAATGAACTT | AACTATTCTA | GATCTATATA | ATGGTTCAAA | TCTTTCTCTA | AATTCCTCAA | .9480 |
| 25 | TACTATAGTA | ATTAGTAGTG | ATATCGAGAA | AGAACGCTAA | ATTCTCTAAA | TTGATCATAT | 9540 |
| 25 | TTTTATGAAA | TCTATTTTTA | TACTTCAAGC | TCTCACAAAA | TCCATCCCAG | TCATTATTTG | 9600 |
| | CTACAATTAG | ATTTTTATTT | GTATATTTTT | TATCGTTTAT | GATTTTAGCG | CCTACTAAAT | 9660 |
| 30 | CTTCCAACAC | TCGTCTATCT | AAATTTTCAT | CATCTTTAAA | AAGTTCATTT | AAAATACAAC | 9720 |
| | TTATTTGAGC | TTCCTCAACA | TTAAATATAC | TCCAGTCGTC | TTTTAATGCT | ATTTCAATCT | 9780 |
| | TTTTACCTTC | TTTTGGGCTA | AAAGTATCTG | GTAAATTTAT | ACTAATATCA | TATAATTCTA | 9840 |
| 35 | | | | | | GTCGTATCAA | 9900 |
| | | • | | • | | CTATTTATCA | 9960 |
| | • | | | | | TTTCATTTTG | 10020 |
| 40 | | | | | | TTTTAATTAA | 10080 |
| | ACCTAATCCT | GGATaCTTAT | TATTTTCATT | TAATTCTTCA | AATTGTCCCA | AGCGCATAAG | |
| 45 | ATCTATTTTT | AATATCTAAG | TTTTTTGACC | ATGTTACTAA | TT | | 10182 |

(2) INFORMATION FOR SEQ ID NO: 117:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3491 base pairs(B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

55

| | AACTCAGGCA | ATTGAAACAG | CATTAGGTGC | TTCATTACAA | CATGTCATTG | TAGATTCAGA | 60 |
|-----------|----------------|----------------|------------|------------|------------------|--------------------|------|
| | AAAAGATGGA | CGCCAGGCTA | TTCAATTTTT | AAAAGAACGT | AATTTAGGTC | GTGCGACGTT | 120 |
| 5 | TTTACCATTA | AATGTTATAC | AGAGTAGAGT | GGTAGCGACT | GATATTAAAT | CTATTGCTAA | 180 |
| | AGAGGCAAAC | GGATTTATTA | GTATCGCTTC | GGAAGCAGTT | AAAGTAGCAC | CAGAATATCA | 240 |
| o | AAATATTATC | GGGAATTTAT | TAGGTAATAC | GATTATCGTT | GATCATTTAA | AGCATGCAAA | 300 |
| U | TGAATTGGCA | CGTGCGATTA | AATATCGAAC | TCGTATTGTT | ACTTTGGAAG | GTGATATTGT | 360 |
| | AAATCCTGGT | GGLTCTATGA | CTGGTGGTGG | CGCTCGTAAG | TCAAAAAGTA | TTCTGTCTCA | 420 |
| 5 | AAAAGACGAG | TTGACAACAA | TGAGACACCA | ATTAGAAGAT | TACTTGCGTC | AAACAGAATC | 480 |
| | ATTTGAACAA | CAATTTAAAG | AGTTGAAGAT | AAAAAGTGAT | CAATTAAGTG | AACTGTATTT | 540 |
| | TGAAAAAAGT | CAAAAGCATA | ATACACTTAA | AGAGCAAGTG | CATCATTTTG | AAATGGAGCT | 600 |
| 20 | CGATAGATTA | ACTACACAAG | AAACACAAAT | AAAAAATGAT | CATGAAGAAT | TCGAATTTGA | 660 |
| | AAAAAATGAT | GGTTATACGA | GTGACAAAAG | TCGACAAACT | TTGAGTGAAA | AAGAAACTTA | 720 |
| | TCTAGAAAGT | ATTAAAGCAT | CTTTAAAACG | ACTAGAAGAT | GAAATTGAAC | GCTACACAAA | 780 |
| ?5 | ACTTTCTAAA | GAAGGTAAGG | AAAGCGTTAC | TAAAACACAA | CAAACCTTAC | ATCAGAAACA | 840 |
| | ATCTGATCTT | GCTGTGGTTA | AAGAGCGTAT | TAAAACACAA | CAACAGACAA | TAGATCGATT | 900 |
| | AAATAATCAA | AATCAACAAA | CTAAACATCA | ATTAAAAGAT | GTTAAAGAAA | AAATTGCATT | 960 |
| 30 | CTTTAATTCG | GATGAAGTGA | TGGGCGAACA | AGCTTTTCAA | AATATTAAAG | ATCAAATTAA | 1020 |
| ٠ | TGGTCAACAA | GAAACGAGAA | CACGCTTATC | AGATGAATTA | GATAAATTGA | AACAACAACG | 1080 |
| 35 | TATTGAGTTG | AATGAACAAA | TCGATGCGCA | AGAAGCTAAA | CTACAAGTTT | GTCACCAAGA | 1140 |
| | TATTTTAGCT | ATCGAAAATC | ACTACCAAGA | TATTAAAGCT | GAACAATCAA | AGCTAGATGT | 1200 |
| | ATTÄÄTTCAT | CATGCGATAG | ATCATTAAAT | GATGRATATC | AATTGACTGT | TGAACGTGCG | 1260 |
| 10 | AFATCTGAAT | ATACGAGTGA | TGrATCGATg | ACGCATTACG | TAAAAAAGTT | AAGTTAATGr | 1320 |
| | AGaTGyCGAT | TGATGrACTA | GGTCCTGTAA | ACTTAAATGC | AATTGAACAA | TTTGAAGAGT | 1380 |
| | TAAATGAACG | TTATACATTT | TTAAGTGAAC | AACGTACAGA | TCTTCGTAAA | GCTAAAGAAA | 1440 |
| 45 | CATTAGAGCA | AATTATAAGT | GAAATGGATC | AAGAGGTTAC | TGAAAGATTT | AAAGAAACTT | 1500 |
| | TCCATGCTAT | TCAAGGACAT | TTTACAGCTG | TGTTCAAACA | ATTGTTTGGT | GGAGGCGATG | 1560 |
| | CAGAATTGCA | ATTAACTGAA | GCCGATTATT | TAACAGCTGG | TATTGATATT | GTGGtACAAC | 1620 |
| 50 | CACCGGGTAA | AAAGTTGCAA | CATTTATCGT | TACTGAGTGG | TGGTGAGCGT | GCATTAACTG | 168 |
| | ~px 4~r(~r~r~r | a cora overcea | ATTTT | таасатстес | ACCTEMENT OF THE | ል ተልተተልርቍተር | 174 |

| | TATCAGACGA | AACACAATTC | ATTGTTATTA | CACACCGTAA | AGGAACAATG | GAATTTGCAG | 1860 |
|----|------------|------------|------------|------------|------------|------------|------|
| | ATAGGTTATA | CGGTGTAACA | ATGCAAGAAT | CAGGTGTTAC | TAAACTTGTG | AGTGTGAATT | 1920 |
| 5 | TAAATACAAT | AGATGATGTG | TTGAAGGAGG | AGCAATAATG | AGCTTTTTTA | AACGCTTAAA | 1980 |
| | AGATAAGTTT | GCAACAAATA | AAGAAAATGA | AGAAGTTAAA | TCCTTAACAG | AAGAACAAGG | 2040 |
| 10 | TCAAGACAAA | TTAGAAGATA | CACATTCTGA | AGGTTCAACG | CAGGACGCAA | ATGATTTAGC | 2100 |
| ,, | AGAAAATGCT | GAAGTGAAAA | AGAAGCCACG | CAAGTTGAGT | GAAGCGGATT | TTGATGACGA | 2160 |
| | TGGCTTAATA | TCAATTGAAG | ATTTTGAAGA | AATTGAAGCT | CAAAAAATGG | GTGCTAAATT | 2220 |
| 15 | TAAAGCAGGA | CTCGAAAAAT | CTCGTCAAAĀ | TTTCCAAGAA | CAATTAAATA | ATTTGATAGC | 2280 |
| | GAGATATCGT | AAAGTAGATG | AAGACTTTTT | TGAAGCTTTA | GAAGAAATGT | TAATCACTGC | 2340 |
| | AGACGTCGGT | TTTAATACAG | TGATGACGTT | AACTGAAGAA | TTACGTATGG | AAGCACAACG | 2400 |
| 20 | ACGTAATATT | CAAGATACTG | AAGATTTGCG | TGAAGTCATT | GTTGAAAAGA | TCGTAGAGAT | 2460 |
| | TTACCATCAA | GAAGATKATA | ATTCAGAAGC | TATGAĄCTTA | GAAGATGGTC | GTTTAAATGT | 2520 |
| | CATTTTAATG | GTTGGTGTGA | ATGGTGTTGG | TAAAACAACA | ACAATTGGAA | AATTAGCTTA | 2580 |
| 25 | CCGATATAAA | ATGGAAGGTA | AAAAAGTAAT | GTTAGCTGCG | GGCGATACTT | TTAGAGCGGG | 2640 |
| | TGCTATTGAT | CAATTGAAAG | TTTGGGGCGA | ACGTGTTGGT | GTAGACGTAA | TTAGCCAAAG | 2700 |
| 30 | TGAAGGTTCT | GATCCAGCTG | CTGTTATGTA | TGATGCGATT | AATGCCGCTA | AAAACAAAGG | 2760 |
| | TGTTGATATT | TTAATCTGTG | ATACCGCTGG | ACGTTTACAA | AATAAmaCAA | ATCTAATGCm | 2820 |
| | AGAATTAGAA | AAAGTTAAGC | GTGTAATTAA | TCGAGCAGTG | CCAGATGCGC | CTCATGAAGC | 2880 |
| 35 | ATTACTATGT | TTAGATGCTA | CAACTGGTCA | GAATGCGTTG | TCACAAGCTA | GAAACTTTAA | 2940 |
| | AGAAGTAACA | AATGTTACAG | GTATTGTATT | AACGAAATTA | GATGGTACAG | CCAAAGGTGG | 3000 |
| | TATCGTATTA | GCCATTCGTA | ATGAATTGCA | CATCCCAGTT | AAATATGTAG | GTTTAGGTGA | 3060 |
| 40 | GCAATTAGAT | GACTTACAAC | CATTTAACCC | TGAAAGTTAT | GTCTACGGCT | TATTCGCTGA | 3120 |
| | TATGATTGAA | CAAAATGAAG | AAATAACAAC | AGTTGAAAAT | GATCAAATTG | TAACAGAAGA | 3180 |
| | AAAGGACGAT | AATCATGGGT | CAAAATGATT | TAGTLAAAAC | GTTACGAATG | AATTATTTGT | 3240 |
| 45 | TTGATTTTaT | CAATCCTTAT | TGACGAATAA | ACAACGTaAT | TATTTGGAAT | TATTTTATCT | 3300 |
| | TGAAGATTAT | TCTTTAAGTG | AAATCGCAGa | TACTTTTAAT | GTGAGTAGAC | AAGCAGTTTA | 3360 |
| 50 | TGATAATATA | AGAAGAACTG | GCGATTTAGT | TGAAGATTAT | GAAAAGAAAT | TGGAATTATA | 3420 |
| - | CCAGAAATTT | GAGCAACGCC | GAGAAATATA | TGATGAAATG | AAACCACATT | TAAGTAATCC | 3480 |
| | AGAACAAATA | C | | | | | 3491 |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4253 base pairs
- (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 118:

| 10 | | | | _ | | | |
|-----|------------|------------|------------|------------|------------|------------|------|
| ,,, | AGTACGTTTT | ATAATTATAA | GTACGTAATT | AACATATTAA | CATATCGCAA | GTATGTATTT | 60 |
| | AAATAAgATT | GTTATAATTT | CAAAGTTCAT | CCAAGaTTAT | GGCGTTTGCA | TTTACCTATT | 120 |
| 15 | AAAAACGTTA | TTATATCAAA | GATGCGAAAG | ATAATACGGG | TTTATTTTAT | GAAAGTGAGA | 180 |
| | AGGATAAAAT | GGATAATGAG | CAACGCTTAA | AAAGAAGAGA | GAATATAAGG | AATTTCTCGA | 240 |
| | TTATAGCACA | TATTGACCAC | GGAAAATCTA | CATTGGCTGA | TAGAATTTTA | GAAAATACCA | 300 |
| 20 | AATCAGTTGA | AACAAGAGAT | ATGCAAGATC | AGTTACTAGA | TTCAATGGAT | TTAGAAAGAG | 360 |
| | AACGTGGTAT | TACAATCAAA | TTAAACGCgT | ACGTTTAAAG | TACGAAGCTA | AAGATGGAAA | 420 |
| | TACTTATACA | TTCCATTTAA | TCGATACGCC | TGGACACGTC | GATTTTACAT | ATGAAGTGTC | 480 |
| 25 | ACGTTCTTTG | GCAGCTTGTG | AGGGCGCGAT | TTTAGTAGTA | GATGCGGCTC | AAGGTATCGA | 540 |
| | AGCACAAACA | TTAGCAAATG | TTTATTTAGC | ATTAGATAAT | GAGTTAGAGT | TATTGCCTGT | 600 |
| 30 | TATTAACAAA | ATTGATTTAC | CTGCTGCAGA | ACCTGAACGC | GTGAAACAAG | AAATTGAAGA | 660 |
| 30 | TATGATAGGT | TTAGACCAAG | ACGATGTTGT | TTTAGCAAGT | GCTAAATCTA | ACATTGGAAT | 720 |
| | TGAAGAGATA | CTAGAGAAAA | TAGTTGAAGT | TGTGCCAGCT | CCAGATGGTG | ACCCAGAAGC | 780 |
| 35 | ACCACTAAAA | GCGTTAATAT | TTGATTCTGA | GTATGATCCA | TATAGAGGGG | TAATTTCATC | 840 |
| | GATAAGAATT | GTGGACGGTG | TTGTTAAAGC | CGGAGATAAA | ATTCGAATGA | TGGCCACTGG | 900 |
| | TAAAGAGTTC | GAAGTAACAG | AAGTTGGAAT | TAATACACCT | AAGCAGCTTC | CAGTTGATGA | 960 |
| 40 | ATTÄACAGTT | GGTGATGTTG | GTTATATTAT | TGCAAGTATT | AAAAATGTTG | ATGATTCTAG | 1020 |
| | GGTTGGTGAC | ACCATCACAT | TAGCTAGTAG | ACCTGCATCA | GAACCATTGC | AAGGTTATAA | 1080 |
| | GAAAATGAAT | CCAATGGTAT | ATTGCGGACT | GTTCCCAATA | GATAACAAAA | ATTATAATGA | 1140 |
| 45 | TTTAAGAGAA | GCATTAGAAA | AATTACAATT | GAATGATGCA | TCATTAGAAT | TTGAGCCTGA | 1200 |
| | ATCGTCACAA | GCATTAGGTT | TTGGTTATAG | AACTGGTTTC | TTAGGTATGT | TACACATGGA | 1260 |
| | AATAATTCAA | GAAAGAATTG | AAAGAGAATT | TGGTATTGAA | TTAATTGCAA | CTGCACCATC | 1320 |
| 50 | TGTAATTTAT | CAATGTGTTT | TAAGGGACGG | TTCAGAAGTG | ACGGTTGATA | ACCCAGCACA | 1380 |
| | AATGCCAGAT | CGTGATAAAA | TTGATAAAAT | ATTTGAGCCA | TATGTTCGTG | CAACTATGAT | 1440 |

55

| | TATAAATATG | GACTATTTAG | ATGATATTCG | TGTAAATATT | GTTTATGAAT | TACCTTTAGC | 1560 |
|----|------------|------------|------------|------------|------------|------------|------|
| _ | TGAAGTTGTA | TTTGATTTCT | TCGATCAACT | TAAATCTAAT | ACTAAAGGAT | ATGCATCATT | 1620 |
| 5 | TGATTATGAA | TTCATCGAAA | ATAAAGAAAG | TAATTTAGTC | AAGATGGATA | TTTTATTAAA | 1680 |
| * | TGGTGATAAA | GTGGATGCGC | TAAGCTTCAT | AGTTCATAGA | GATTTTGCAT | ATGAACGTGG | 1740 |
| 10 | TAAAGCATTA | GTTGAAAAAC | TTAAAACGTT | AATTCCAAGA | CAGCAATTTG | AAGTACCTGT | 1800 |
| | ACAGGCTGCA | ATAGGACAAA | AAATTGTAGC | GCGTACAAAT | ATTAAATCAA | TGGGTAAAAA | 1860 |
| • | CGTTTTAGCT | AAATGTTATG | GCGGTGACAT | AAGCCGTAAA | CGTAAATTAC | TTGAAAAACA | 1920 |
| 15 | AAAAGCAGGT | AAAGCTAAGA | TGAAAGCAGT | TGGTAATGTT | GAAATTCCAC | AAGATGCTTT | 1980 |
| , | CTTGGCTGTA | TTGAAAATGG | ATGATGAATA | ATTTTAAAAA | ATCAATTAAC | AATTTACAAT | 2040 |
| | GAATAAAGTT | TAATAACTAA | AAAGAGGGAG | CCTAGGATAA | ATTAACGTCC | TGGGCTTTAC | 2100 |
| 20 | AATGTTATAT | TGGCAGCCAT | CGACAGAGTT | AAAATGAGCT | TATAACAATG | GGGCCCCAAC | 2160 |
| | ACAGAAGCTG | ACGAAAAGTC | AGCTTACTAT | AATGTGCAAG | TTGGGGTGGG | GCCCCAACAT | 2220 |
| ae | AGAGAATTTC | GAAAAGAAAT | TCTACAGGCA | ATGCAAGTTG | GGGTGGGACG | ACGAAATÁAA | 2280 |
| 25 | TTTTGCGAAA | ATATCATTTC | TGTCCCACTC | CCTTATGCAT | GAGTTTTACT | CATGTAATTT | 2340 |
| | TATTTTTAAG | GACATATTAC | ATCTGGCTAA | TGTGTAAGAG | CCACTACATA | ATAAATCATT | 2400 |
| 30 | AGTGGTTCTT | TATTATTTCT | ATCTCACTCC | CTCTAAACAA | GAATAAATAT | TAAAATGAAT | 2460 |
| | CGATATATTA | GACAATCATT | GATTAAACGT | TAAAGTTAAA | AGTAAGAATA | ATTGCAGATA | 2520 |
| | GTCCAACAGG | ATATAGCCGA | TTGGATAAAA | AGTCTGAGAA | GCGGGGCATT | AAAATGACGG | 2580 |
| 35 | TACAAAGTGC | ATATATACAT | ATTCCATTTT | GTGTAAGAAT | ATGTACATAT | TGTGATTTCA | 2640 |
| - | ATAAATATTT | TATACAGAAT | CAACCTGTAG | ATGAGTACTT | AGATGCACTA | ATCACAGAAA | 2700 |
| | TGTCTACAGC | AAAATATAGG | ATCTTAAAGA | CCATGTATGT | AGGTGGCGGC | ACACCAACGG | 2760 |
| 40 | CCCTTTCTAT | TAATCaGTTG | GAAAGATTAC | TTAAAGCAAT | ACGTGATACG | TTTACAATCA | 2820 |
| | CAGGCGAGTA | TACATTTGAA | GCAAATCCTG | ATGAGTTAAC | TAAAGAGAAA | GTCCAACTAT | 2880 |
| 45 | TAGAGAAATA | TGGAGTAAAA | AGGATTTCAA | TGGGCGTTCA | AACATTCAAG | CCGGAGTTAT | 2940 |
| | TGTCTGTTTT | AGGTAGAACG | CACAATACTG | AAGATATTTA | CACTTCGGTG | TTAAATGCTA | 3000 |
| | AAAACGCAGG | TATTAAATCA | ATCAGTTTAG | ATTTAATGTA | TCATTTACCG | AAACAGACGA | 3060 |
| 50 | TTGAAGATTT | TGAACAAAGT | TTAGATCTAG | CTTTAGATAT | GGATATTCAA | CATATTTCGA | 3120 |
| | GTTACGGCTT | AATACTTGAA | CCTAAAACCC | AATTTTATAA | TATGTATAGA | AAAGGCTTGC | 3180 |
| | TCAAACTTGC | TAATGAGGAT | TTAGGTGCTG | ACATGTATCA | GTTGCTGATG | TCTAAGATAG | 3240 |

| AACATAATAA | GGTTTACTGG | TTTAATGAGG | AATATTATGG | ATTTGGAGCA | GGTGCAAGTG | 3360 |
|------------|--------------|-------------|------------|------------|------------|------|
| GTTATGTAGA | TGGTGTGCGT | TATACGAATA | TCAATCCAGT | GAATCATTAT | ATCAAAGCTA | 3420 |
| TAAATAAAGA | AAGTAAAGCA | ATTTTAGTAT | CAAATAAACC | TTCTTTGACT | GAGAGAATGG | 3480 |
| AAGAAGAAAT | GTTTCTTGGG | TTGCGTTTAA | ATGAAGGTGT | GAGTAGTAGT | AGGTTCAAAA | 3540 |
| AGAAGTTTGA | CCAATCTATT | GAAAGTGTCT | TTGGTCAAAC | AATAAATAAT | TTAAAAGAGA | 3600 |
| AGGAATTAAT | TGTAGAAAAG | AACGATGTGA | TTGCACTTAC | AAATAGAGGG | AAAGTCATAG | 3660 |
| GTAATGAGGT | TTTTGAAGCT | ттсстаатаа | ATGATTAAAA | AAAATTGAAA | TTTCGAGTCT | 3720 |
| TTAACATTGA | CTTACTTTGA | CCAATTTGAT | AAATTATAAT | TAGCACTTGA | GATAAGTGAG | 3780 |
| TGCTAATGAG | GTGAAAACAT | GATTACAGAT | AGGCAATTGA | GTATATTAAA | CGCAATTGTT | 3840 |
| GAGGATTATG | TTGATTTTGG | ACAACCCGTT | GGTTCTAAAA | CACTAATTGA | GCGACATAAC | 3900 |
| TTGAATGTTA | GTCCTGCTAC | AATTAGAAAT | GAGATGAAAC | AGCTTGAAGA | TTTAAACTAT | 3960 |
| ATCGAGAAGA | CACATAGTTC | TTCAGGGCGT | TCGCCATCAC | AATTAGGTTT | TAGGTATTAT | 4020 |
| GTCAATCGTT | TACTTGAACA | AACATCTCAT | CAAAAAACAA | ATAAATTAAG | ACGATTAAAT | 4080 |
| CAATTGTTAG | TTGAGAATCA | ATATGATGTA | TCATCAGCAT | TGACATATTT | TGCAGATGAA | 4140 |
| TTATCAAATA | TATCTCAATA | TACAACTTTA | GTTGTTCATC | CTAATCATAA | ACAAGATATT | 4200 |
| ATCAATAATG | TACACTTGAT | TCGTGCTAAT | CCTAATTTAG | TTATAATGGT | TAT | 4253 |
| (2) INFORM | ATION FOR SE | Q ID NO: 11 | L9: | | | |
| (i) SF | COUENCE CHAP | 20075819710 | ₹. | | | |

- (A) LENGTH: 3395 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

40

5

10

15

20

25

30

35

45

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 119:

TCCCTAATCG AACAAATTA TGCGCATAAA CAAAGTAGAT TGATATAAAA TTCTTAATTA 60 TCAGAATATA TTTACAAATC TGAATTTTAT TAGTATATTG GrTAGTrTTC ATAGAGGCAT 120 GACGGTaTTT GAGCAGGATT TTAAATCGGg ATTTTATAAT CGATTTAAGA GAGGCCACtT 180 TGCTTGCACA TTAATACTGT CAATGGGAGG GGAATGTATA TGAGTTAAGC ACATCAATTA 240 ATTCAAGAGG ATGAACATTA TTTTGCGAAA TCAGGACGTA TTAAATATTA TCCGTTAGTG 300 ATTGATCATG GATATGGAGC AACATTGGTT GATATTGAGG GGAAGACATA TATCGATTTG 360 TTATCGAGTG CGAGTTCTCA AAACGTAGGT CATGCACCTA GAGAAGTAAC AGAAGCGATA 420

| | GTACGTTTAG | CTAAGAAGCT | TTGTGAGATT | GCACCTGGAG | ATTTTGAAAA | AAGAGTGACC | 540 |
|------------|------------|------------|------------|------------|------------|------------|------|
| _ | TTCGGATTAA | CCGGATCAGA | CGCAAATGAT | GGCATCATTA | AATTTGCCAG | AGCATATACA | 600 |
| 5 | GGGCGTCCTT | ATATCATTAG | TTTCACTAAT | GCATATCATG | GTTCAACTTT | TGGCTCATTG | 660 |
| | TCTATGTCAG | CTATTAGTTT | AAATATGCGC | AAACATTATG | GTCCGTTATT | GAATGGTTTT | 720 |
| 10 | TATCATATTC | CGTTTCCAGA | TAAATATCGT | GGTATGTACG | AGCAGCCACA | AGCTAATTCA | 780 |
| | GTAGAAGAAT | ATTTAGCACC | CTTAAAAGAA | ATGTTTGCGA | AGTATGTACC | TGCTGACGAA | 840 |
| | GTAGCATGTA | TTGTTATTGA | AACGATACAA | GGCGATGGTG | GACTTTTAGA | ACCAGTTCCA | 900 |
| 15 | GGGTATTTTG | AAGCGTTAGA | AAAGATTTGT | CGTGAACATG | GTATTTTAAT | CGCTGTCGAT | 960 |
| | GATATTCAAC | AAGGTTTTGG | GAGAACAGGT | ACATGGAGTT | CAGTCTCGCA | TTTTAATTTT | 1020 |
| | ACGCCTGATT | TAATCACTTT | CGGAAAATCC | TTAGCAGGTG | GTATGCCTAT | GTCAGCAATT | 1080 |
| 20 | GTTGGACGCA | AAGAGATTAT | GAATTGTTTA | GAAGCACCAG | CACATTTATT | TACAACAGGT | 1140 |
| | GCTAATCCAG | TTAGTTGTGA | AGCTGCATTA | GCCACAATTC | AAATGATTGA | AGATCAGTCG | 1200 |
| | CTTCTTCAGG | CTAGTGCGGA | AAAAGGGGAA | TATGTTAGGA | AACGAATGGA | TCAATGGGTA | 1260 |
| 25 | TCTAAATACA | ATAGTGTAGG | CGATGTTAGA | GGTAAAGGTC | TGAGCATTGG | TATTGATATT | 1320 |
| | GTTTCCGACA | AAAAACTCAA | AACACGTGAT | GCCAGTGCGG | CACTTAAAAT | TTGTAATTAC | 1380 |
| 3 <i>0</i> | TGCTTTGAGC | ATGGCGTAGT | TATTATAGCT | GTAGCAGGAA | ATGTGTTGCG | ATTCCAACCG | 1440 |
| | CCATTGGTAA | TAACATATGA | GCAATTAGAC | ACGGCGTTAA | ACACTATAGA | AGATGCACTG | 1500 |
| | ACTGCTTTGG | AAGCAGGTAA | CTTAGATCAA | TATGACATAT | CTGGACAAGG | TTGGTAATAG | 1560 |
| 35 | CGATTATCTT | AATATAAAAT | AAAAAATCAT | TTCCACATCT | GGATGTTAAT | CAGATGGGAA | 1620 |
| • | ATGATTTTTT | TTATTTTTTA | TTTTGGTGGG | TGGTATTCAG | CTACGTCATT | TTTCTTAGAA | 1680 |
| | TGTGTAAGTC | CATAACTTAA | ATATAGGATG | ATACCAACAA | TAAACCAAAT | TAAAGTGTAT | 1740 |
| 40 | AATTTCGCTT | CGAATCCTAA | TCCCCAGAAT | ACTAGCAATA | CTAAAACAAA | TGTAATTGCT | 1800 |
| | GGTAACACAG | GATATAAAGG | TAATTTAAAT | GCAGGAATTG | GTAGATCTTT | ACCTTCACGC | 1860 |
| | TTTCTCAAAC | GATACATTGC | TAATGAAACG | AACATAAATG | CAACAAGTGT | ACCTGCTGAA | 1920 |
| 15 | ATTAATTGTG | CTAAAAATGC | GAATGGGAAC | ATAGAACCAA | TTAAAACACC | AATAATAGTA | 1980 |
| | AGTATAACTA | GTGCGCGATT | AGGTAAATGT | TTGTCGTTTA | AGTGGCTTAA | CCATGAAGGT | 2040 |
| 5 0 | AATAAGCCGT | CACGTCCAAA | TGAATAAAGT | AAACGTGAGC | CTGCTAACAT | CATACCAATT | 2100 |
| · · | AATGCTGTAA | ACATACCGAT | AACAGAGATA | GCTTGAACAA | TAGCTGCTAC | AACACCATGA | 2160 |
| | CCACTTTGAC | GTAAAGCCCA | ACCAACAGGT | TCAGCATTGT | TTGCGTATTG | TGAGTAATGG | 2220 |

670

| | CCAAGAATAC | CTCTAGGCAT | TGTCTTTTGA | GGATCAAGTG | CTTCTGCTGA | GTTTGCTGCG | 2340 |
|----|------------|------------|------------|------------|------------|------------|------|
| | ATAGAATCGA | AACCGATATA | CGCTAAGAAA | ATCATTGAAA | CACCAGCATA | TATGCCTTGC | 2400 |
| 5 | CATCCACCAA | AGTCACCTGT | AGCAGTTACT | TTGTGTTCTG | GAATAAATGG | CACATAGTTA | 2460 |
| | CTAACATTTA | TTGCTGTTAA | ACCTACGATG | ACAAATAAAA | TAATAGCTAA | TACTTTTAAA | 2520 |
| 10 | ATAACTAAAA | TATTTTCCAT | ACGAGCTGCT | TCCGACATAC | CACGTGATAG | TAATAATGCA | 2580 |
| 10 | GTTAATAAAA | TAACGATAGC | AGCAATAATA | TCGATAAAAC | CGCCATTTGT | ACCAAATGGA | 2640 |
| | TTTGATAATG | CTGCAGGTAA | TTCGATGCCA | ATTGGTTTCA | CAAGTCCGCG | TAAATTCGCT | 2700 |
| 15 | GAGAATCCTG | ATGCAACAAA | GGCTACGGCG | ATAAAAATAT | CAGCTAATAG | AGCCCAACCG | 2760 |
| | GCAACCCATC | CAAAAAATTC | ACCAAATAAT | ACATTGACCC | AAGAATAGGC | TGAACCTGCA | 2820 |
| | AATGGCATAG | CGGCAGCCAT | TTCTGCATAA | GTAAATGCAA | CTAAACCAGC | AACAATAGCA | 2880 |
| 20 | GCGAGTAAGA | ATGATAACGC | AACGGCCGGT | CCTGCATGTT | CTGCAGCAAC | AATGCCAGGT | 2940 |
| | AGCGTAAAGA | TAGATGTCGA | TACAATTGTT | CCTACACCTA | AAGCTAAGAA | ATCACGCACC | 3000 |
| | CGAAGTGTAC | GCTTTAAATG | ACCATCTTTA | TTTTGATAGA | TAGCCGGATC | CTCTTTTCGT | 3060 |
| 25 | GCTATTTTAT | TGAAAAAACT | TCCCATAAAC | TTTCCTCCCA | AACATTCATA | AACAATTCTA | 3120 |
| | TACGGTGTTT | TTTAATATGT | TATATCATAG | CACAAATAAT | CAATATTTTG | TCTAAAAATT | 3180 |
| • | CTGAAAAATC | ACAACTTTAT | GTTACGTATT | AATGACTTGT | CTTGATAACA | TCCATAGATT | 3240 |
| 30 | TTTTAAATGA | TAAAACTGAT | TATAACAGAT | ATTAAATGAA | TAAGTACTAT | TTTTTGCnAA | 3300 |
| | TTTTCTAACA | ATTTTGCACA | TTATATGTTT | AAAATCAATT | TCATGTTTAT | GGTCTGATTG | 3360 |
| 35 | GCTAGTGTGT | atgaaatgta | Antctttgac | TnnGA | | | 3395 |

(2) INFORMATION FOR SEQ ID NO: 120:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13508 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

45

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 120:

ATCAGGTAAT GCCATGCGTT TAGCTGAAAA TTTTTTCAGA ACGTTTAAGT GATATCGGAC 60 ATCAAGTTGT TTTGATGTCA ATGGATGAAT ATGATACGAC AAACATCGCG CAGTTAGAAG 120 ATTTATTTAT TATTACGTCT ACTCATGGTG AAGGAGAACC GCCTGATAAT GCATGGGATT 180 TCTTTGAATT TTTAGAAGAC GATAACGCAC CTAATTTAAA TCATGTGAGA TATTCAGTAC 240

55

| | TACTAGAAAA | TCTAGGCGCT | GAGCGTATAT | GTAAGCGTGT | AGATTGTGAT | ATTGATTATG | 360 |
|----------|-------------|------------|------------|------------|------------|------------|------|
| | AAGAAGACGC | AGAAAAGTGG | ATGGCAGACA | TCATTAATAT | TATTGATACC | ACATCAGAAG | 420 |
| 5 | GTATTCAAAG | TGAATCGGTG | ATAAGTGAAT | CAATTAAGTC | TGCCAAAGAA | AAGAAATATT | 480 |
| | CTAAATCAAA | TCCATACCAA | GCAGAAGTAT | TAGCGAATAT | CAATTTAAAT | GGTACCGATT | 540 |
| 10 | CAAATAAAGA | AACACGACAT | ATAGAATTTT | TACTTGATGA | TTTTAGTGAA | TCATATGAAC | 600 |
| 70 | CAGGAGATTG | TATAGTAGCA | TTACCGCAAA | ACGACCCTGA | ATTGGTTGAA | AAACTAATAT | 660 |
| | CCATGTTAGG | TTGGGATCCG | CAATCTCCGG | TGCCAATTAA | TGATCATGGT | GATACAGTTC | 720 |
| 15 | CTATTGTTGA | AGCACTAACA | TCACATTTTG | AATTTACTAA | ATTAACATTG | CCATTATTGA | 780 |
| | AAAATGCAGA | TATCTATTTT | GACAATGAAG | AATTATCTGA | ACGTATTCAA | GATGAGTCAT | 840 |
| | GGGCGCGTGA | ATATGTTATA | AATCGGGACT | TTATAGATTT | AATAACAGAT | TTTCCAACTA | 900 |
| 20 | TAGAATTACA | ACCTGAGAAT | ATGTATCAAA | TCCTTAGAAA | ATTACCACCA | AGAGAGTATT | 960 |
| | CGATTTCTAG | TAGTTTTATG | GCAACGCcAG | ATGAAGTGCA | TATTACCGTT | GGTACGGTTC | 1020 |
| | GTTATCAAGC | ACATGGACGT | GAGAGAAAAG | GTGTATGCTC | GGTTCATTTT | GCTGAGCGAA | 1080 |
| 25 | TTAAACCAGG | CGATATAGTA | CCAATTTATT | TGAAGAAAA | TCCGAACTTC | AAATTTCCGA | 1140 |
| | TGAAGCAAGA | TATACCGGTT | ATTATGATTG | GACCAGGTAC | TGrAATTGCT | CCTTTTAGAG | 1200 |
| 20 | CATATTTACA | AGAACGTGAA | GAACTTGGTA | TGACTGGAAA | AACATGGTTG | TTCTTTGGTG | 1260 |
| 30 | ATCAACACCG | TAGTTCTGAC | TTTTTATATG | AAGAAGAAAT | AGAAGAATGG | CTTGAAAATG | 1320 |
| | GAAACTTAAC | ACGCGTAGAT | TTAGCATTTT | CAAGAGACCA | AGAACACAAA | GAATATGTAC | 1380 |
| 35 | AGCATCGTAT | AATGGAAGAA | AGTAAACGTT | TCAATGAATG | GATTGAGCAA | GGCGCACAAT | 1440 |
| | CTATATTTGT | GGCGATGAAA | AATGTATGGC | GAAAGATGTC | CATCAAGCCA | TTAAAGATGT | 1500 |
| | ATTGGTAAAA | GAACGTCATA | TTTCTCAAGA | AGAAGCAGAG | TTATTATTGC | GACAAATGAA | 1560 |
| 40 | ACAACAACAA | CGCTATCAAC | GTGATGTTTA | TTAGCGATTG | GTGTTAAATA | TTTTAAGGTG | 1620 |
| | TAATGATGTA | AAAAGATATA | AAGGATGTTG | CTCAACATGA | ATATGCCATT | AATGATAGAT | 1680 |
| | TTAACAAATA | AAAATGTCGT | CATAGTTGGT | GGAGGCGTCG | TTGCAAGTCG | TCGGGCACAA | 1740 |
| 45 | ACATTAAATC | AATACGTTGA | ACATATGACG | GTCATCAGTC | CGACAATCAC | TGAAAAACTT | 1800 |
| | CAAAATATGG | TAGATAACGG | TGTCGTCATA | TGGAAAGAAA | AAGAATTTGA | ACCAAGCGAT | 1860 |
| 50 | ATTGTAGACG | CGTATCTAGT | TATTGCAGCA | ACCAATGAGC | CACGTGTCAA | TGAAGCGGTA | 1920 |
| 50 | AAAAAAGCCT | TACCTGAGCA | TGCCCTTTTT | AATAATGTTG | GAGATGCATC | AAATGGCAAT | 1980 |
| | CHALT TALLO | САВСТССАСТ | ACACCGCGAC | аасстааста | телететате | AACTCATCCT | 2040 |

| | TACAGTTCGT | ATATCGACTT | TTTATATACT | TGCCGACAGA | AAATAAAAGT | ACTTGATATA | 2160 |
|------------|------------|------------|-------------------|------------|------------|------------|--------|
| | ACATATAACG | AAAAGCAACA | GTTACTGTCA | CAAATTGTGT | CACAAGAATA | TTTAAATCAT | 2220 |
| 5 | GACAAACAAG | CTCAATTTTT | AGCGTGGTTG | GATGTAAGAT | AATAATAGCG | GACCGTCTAA | 2280 |
| | CCGTCTAAGG | TAAGTCTTCT | TATTTTAACT | TTAACGCTTA | ATCATTGAAA | TTAAGACATG | 2340 |
| 10 | GGCGGCTTTG | TGAATAGTCT | AATAATGAAG | GATTTAAGCG | ATAATGATAT | GCGTTTTAAA | 2400 |
| 10 | TATGAATATT | ACAATAGAGA | AAAAGATACG | TAGAACAAAC | TTAATAAAAT | AGGTGGATAA | 2460 |
| | ATTGAAATCT | GGTTGAAGTC | GTTACTATCA | TAGCGACCTT | TAGCCAGATT | TTTTGTGCAA | 2520 |
| 15 | TAGAAAGCAA | TAATAAAAAT | GATAGATCAA | AATGAAATAC | AGGACAGGAT | ATACAAGGAT | 2580 |
| | TAGTCATGCC | ATGTTATCAA | GTAGGAAAAT | CAAACTTCAC | TATTGATAGT | TACGCAAAAA | 2640 |
| | AGATTTTTT | GATAAAATGA | GATAACTTAA | АТАТААААА | TTATATTAAT | TATAATATTT | 2700 |
| 20 | AAGTTAAAGA | GGGGGATTAT | GTAAATTGTA | TTAAAAGTGG | AGGGAGAAAA | TAATATGAAT | 2760 |
| | AGTGATAATA | TGTGGTTAAC | AGTAATGGGG | CTCATTATTA | TTATTTCAAT | TGTAGGTTTA | 2820 |
| | CTCATTGCCA | AAAAGATAAA | TCCAGTTGTA | GGTATGACAA | TCATACCTTG | CTTAGGGGCA | 2880 - |
| ? 5 | ATGATTTTAG | GATATAGTGT | GACAGATTTG | GTTGGATTTT | TTGCTAAAGG | GTTAGATCAA | 2940 |
| | GTCATCAACG | TTGTTATTAT | GTTTATCTTT | GCCATTATTT | TCTTTGGCAT | CATGAACGAT | 3000 |
| | AGTGGTTTAT | TCAAGCCGCT | TGTCAAACGC | TTAATATTAA | TGACACGAGG | CAATGTCGTC | 3060 |
| 30 | ATTGTCTGTG | CAATGACAGC | TTTAATTGGC | ACAATAGCCC | AATTAGATGG | GGCCGGTGCG | 3120 |
| | GTAACATTTT | TGCTTTCTAT | TCCTGCATTA | TTACCTTTAT | ATAAAGCGTT | AAATATGAAT | 3180 |
| 35 | AAATATTTAT | TGATTTTACT | ATTAGCATTA | AGCGCGGCGA | TTATGAACAT | GGTACCTTGG | 3240 🥌 |
| | GGAGGTCCAA | TGGCTCGTGT | AGCTGCAGTG | TTAAAAGCCA | AAAGTGTCAA | TGAATTATGG | 3300 |
| | TATGGATTAA | TACCTATTCA | AATAATAGGT | TTCATTCTTG | TTATGTTGTT | TGCGGTATAT | 3360 |
| 40 | CTTGGATTTA | AAGAACAGAA | ACGTATCAAA | AAAGCAATAG | agagaaatga | ATTACCGCAA | 3420 |
| | ACACAAGATA | TAGATGTACA | TAAATTAGTT | GAAGTATATG | AACGAGATCA | AGATGTAAGG | 3480 |
| | TTTCCTGTAA | AAGGACGTGC | AAGAACAAAA | TCATGGATAA | AATGGGTGAA | TACAGCTTTA | 3540 |
| 45 | ACTTTAGCTG | TTATTCTATC | GATGTTAATA | AATATTGCGC | CACCTGAATT | TGCATTCATG | 3600 |
| | ATAGGTGTTY | CGTTGGCACT | TGTTATTAAT | TTTAAATCAG | TGGATGAACA | AATGGAACGA | 3660 |
| | TTAAGAGCgC | ATGCGCCGAA | TGCATTAATG | ATGGCTGCAG | TGATTATTGC | AGCAGGTATG | 3720 |
| 50 | TTTTTAGGTG | TACTAAATGA | AACCGGTATG | CTTAAAGCGA | TTGCGACCAA | TTTAATCAAA | 3780 |
| | GTGATTCCTG | CAGAAGTAGG | ACCATACTTG | CATATTATTG | TAGGTTTACT | TGGCGTACCA | 3840 |

| | ACAGCAGGGC | AATTTGGTGT | ACCGTCTGTA | TCAACAGCTT | ATTCAATGGT | CATAGGGAAT | 3960 |
|----|------------|------------|------------|------------|-------------------|------------|-------|
| | ATTATAGGTA | CATTTGTCAG | CCCATTTTCA | CCAGCCTTAT | GGTTGGCAAT | TGGTTTAGCA | 4020 |
| 5 | GAGGCAAACA | TGGGCACGTA | TATTAAGTAT | GCATTCTTTT | GGATTTGGGG | ATTCGCTATC | 4080 |
| | GTTATGTTAG | TAATTGCAAT | GTTGATGGGC | ATTGTGACGA | TTTAAGTATG | AAAAAATAGA | 4140 |
| 10 | AACTATGGTC | ACGTTGCAAA | ATGAAATAAT | AGTTGCATAA | ACATGTCGAA | ATGACGGACG | 4200 |
| | AATCTTTAAA | CAATTTTAAA | AATTAATGAA | ATAATTGTGT | AGAAATATGA | ATTTCACTAA | 4260 |
| | ATGTTAATAA | CTTTGTGACG | TTTTAGTTAA | CAGACTAATA | AAAATTTGAA | AATACTATAT | 4320 |
| 15 | ATAGTGGTAT | AACGTAATGA | GTAGACACAA | TATATAGGAA | GAAGGGGTAA | AATGAATCAA | 4380 |
| • | ATCGAAGAAG | CATTAACGGG | TTTGATTTCT | AAAGATCCTG | CTATTGTTAA | CGAAAATGCT | 4440 |
| | AACAAAGATA | GTGATACATT | TTCAACAATG | AGAGATTTAA | CAGCAGGTAT | CGTTTCTAAA | 4500 |
| 20 | TCTTACGCAT | TAAATCATTT | ATTACCAAAG | CACGTTGCAG | ATGCACATCA | AAGAGGGGAC | 4560 |
| | ATACATTTTC | ACGACTTAGA | TTATCATCCA | TTCCAACCGT | TAACTAACTG | TTGTTTAATA | 4620 |
| | GATGCTAAAA | ATATGCTACA | TAATGGATTT | GAAATAGGCA | ACGCGAATGT | AACTTCACCA | 4680 |
| 25 | AAATCAATAC | AAACTGCATC | AGCGCAGCTT | GTACAAATTA | TAGCCAATGT | TTCTAGCAGT | 4740 |
| | CAATATGGTG | GCTGTAcGGT | TGACCGCGTT | GACGAATTAC | LTAGTACATA | TGCACGACcA | 480.0 |
| 20 | TAATGAAGAA | CAACATAGGA | ATATSCGCAA | AGCAATTTGT | CAAAGAATCT | GAAATTGATC | 4860 |
| 30 | GTTATGTTGA | TCAACAAGTC | ACTAAAGACA | TCAATGATGC | GATTGAAAGT | TTAGAATATG | 4920 |
| | AAATTAATAC | CTTATATACA | TCTAATGGAC | AGACACCTTT | TGTAACATTA | GGATTCGGCT | 4980 |
| 35 | TAGGTACAGA | TCATTTAAGT | CGCAAAATTC | AACAAGCTAT | CTTAAATACT | CGTATCAAAG | 5040 |
| | GCTTAGGAAA | AGACCGCACG | ACAGCGATTT | TCCCGAAACT | TGTATTTTCA | ATTAAAAAAG | 5100 |
| | GAACCAACTT | TAGTCCGCAA | GATCCGAACT | ATGACATTAA | ACAACTAGCA | TTAAAGTGTT | 5160 |
| 40 | CAACGAAACG | TATGTATCCA | GATATTTTAA | ATTATGACAA | ACTCGTAGAA | ATATTAGGTG | 5220 |
| | ATTTCAAAGC | GCCAATGGGT | TGTCGTTCAT | TTTTACCAAG | TTGGAAAGAT | GCGGAAGGTC | 5280 |
| | ATTTTGAAAA | TAATGGTCGT | TGTAATCTTG | GTGTTGTTAC | ACTTAATTTA | CCTAGAATGG | 5340 |
| 45 | CATTAGAATC | TGCCGGTAAT | ATGACGAAAT | TCTGGGAAAT | CTTTTATGAA | CGTATCGATG | 5400 |
| | TGTTACATGA | TGCATTACTT | TATCGTATAA | ATCGTTTGAA | AGATGCTGTA | CCGAATAACG | 5460 |
| | CACCGATTTT | ATATAAAAGT | GGCGCATTTA | ACTATAAATT | AAAAGAAACA | GATGATGTTG | 5520 |
| 50 | CTGAGTTATT | ТАААААТААА | CGTGCAACGA | TTTCAATGGG | CTATATAGGG | TTGTATGAAA | 5580 |
| | CAGCTACTGT | TTTCTATGGT | CCAGACTGGG | AAACATCTCA | AGAAGCAAAA | GCATTTACGC | 5640 |

| | GGTTCAGTAT | TIMCAGTACG | CCGAGTGAAT | CGCTACGGAT | CGTTTTTGTC | GTTTAGACCA | 5760 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | AGAGAGATTT | GGAGATATTA | AAGACATTAC | AGATAAAGGA | TATTATCAAA | ACTCTTTCCA | 5820 |
| 5 | TTATGATGTA | CGTAAAGATG | TTACACCTTT | TGAAAAGTTA | GATTTTGAAA | AAGATTATCC | 5880 |
| | TTATTATGCG | AGTGGTGGTT | TCATTCACTA | TTGTGAGTAT | CCGAAATTGC | AACACAATTT | 5940 |
| 0 | GAAAGCACTA | GAAGCGGTAT | GGGACTACTC | TTATGACAAA | GTTGGTTACT | TAGGTACAAA | 6000 |
| | TATTCCGATT | GATCATTGTT | atgaatgtga | TTACGATGGA | GATTTTGAAG | CAACTGAAAA | 6060 |
| | AGGATTTAAA | TGCCCGAACT | GTGGCAATGA | TAATCCTAAA | ACAGTTGATG | TCGTTAAACG | 6120 |
| 5 | AACATGTGGT | TACCTAGGCA | ATCCAGTTCA | ACGTCCAGTA | ATTAAAGGCC | GTCATAAAGA | 6180 |
| | AATTTGCGCA | CGAGTAAAAC | ATATGAAAGC | GCCTAAAGAA | TGATACTTTT | AGACATTAAA | 6240 |
| | CAAGGACAAG | GTTATATTGC | TAAAATAGAA | TCAAATAGCT | TTGTTGACGG | TGAAGGAGTA | 6300 |
| 20 | AGATGCAGTG | TTTATGTATC | AGGATGTCCA | TTTAATTGTG | TTGGATGTTA | TAACAAAGCC | 6360 |
| | TCACAAAAGT | TCAGATATGG | CGAGAAATAC | ACTGATGAAA | TATTAGCAGA | AATATTAGAT | 6420 |
| | GATTGCGATC | ATGATTATAT | ATCTGGGCTA | AGTCTATTAG | GTGGCGAACC | ATTTTGTAAT | 6480 |
| ?5 | TTGGATATTA | CATTAAATCT | TGTCAAAGCA | TTTCGAGCAC | GTTTTGGAAA | TACAAAGACA | 6540 |
| | ATTTGGGTAT | GGACTGGATT | TTTATATGAA | TATTTAGCAA | ATGATTGTAC | AGAACGTCGA | 6600 |
| | GAGTTATTAT | CATACATTGA | CGTTTTAGTA | GATGGTCTAT | TTATACAACA | CTTATTCAAA | 6660 |
| 30 | CCTGATTTAC | CATATAAAGG | TTCTTTAAAT | CAACGCATTA | TAGATGTACA | ACAATCACTC | 6720 |
| | TCGCATGCGC | GTATGATTGA | ATATATAGTT | AGTTGAATAT | GTATTAGAAG | TCAAGGTAAC | 6780 |
| 35 | ATTCGTTGCC | TTGGCTTCTT | TTTAGGTTAG | GTACATAATT | GAAAGTTAAT | AAAAGCAATT | 6840 |
| | CTTTATAAAA | ATATATTGAT | AGAATATGAC | CTAACAATCA | TTTTGATACC | AATACTAAAA | 6900 |
| | GTTGCATATC | CGTTTTTTAA | AAAAGTTGAA | AGAGAAAAGT | GGTATTTTAG | TGGGAAGGAA | 6960 |
| 10 | GTCTAACTTT | TTGGTAGCGT | TTTACAATAA | ATAAATATTC | GTTAATAACG | TATAAATAT | 7020 |
| | CTTAAATGCC | ATTCTAGTAA | AATTTGTTAA | ATTCGTTAAA | TCGTAACTTA | ACACTGTTAT | 7080 |
| | TTTAGCGCTA | TTAAGGTTTT | GTTTATTACG | GGAAAAATTA | TATAAATATT | CAATAATTGC | 7140 |
| 1 5 | CAAGTTTCAA | attgtatgaa | ATTTGCATTA | TTATTAAATG | TTAGTTATTG | TCAATTTTGT | 7200 |
| | GAATCAATAT | AATTATTACA | TTTTGAGATA | AATCGAAACA | GGATTCATAA | AATTAATAAT | 7260 |
| | TAGGGGGAGC | ACAATTGAAA | AAAGAGAAAG | TTATGGACTG | GACGACCTTT | ATAGGGACAG | 7320 |
| 50 | TAGCTGTACT | TCTTTTTGCA | GTTATACCTA | TGATGGCTTT | TCCAAAAGCA | AGTGAAGATA | 738 |
| | TCATCACTGG | TATTAATAGT | GCCATTTCTG | ATTCAATTGG | TTCGATATAT | TTATTTATGG | 744 |

| | TTGGTAAAGC | AAGTGATAAA | CCAGAATTTA | ATACATTTAC | ATGGGCGGCA | ATGCTGTTTT | 7560 |
|----|------------|------------|--------------|--------------|------------|-------------|-------|
| | GTGCAGGCAT | AGGCTCTGAT | ATTTTATACT | GGGGCGTTAT | TGAATGGGCT | TTTTACTATC | 7620 |
| 5 | AAGTTCCACC | AAATGGCGCG | AAAAGTATGA | GTGATGAAGC | ACTCCAATAT | GCGACGCAAT | 7680 |
| | ATGGTATGTT | CCACTGGGGG | CCAATTGCTT | GGGCTATTTA | TGTTCTACCA | GCATTACCAA | 7740 |
| 10 | TTGGTTATTT | AGTATTTGTT | AAAAAACAAC | CGGTGTATAA | AATTAGTCAA | GCTTGTCGTC | 7800 |
| | CGATTTTAAA | AGGTCAAACA | GATAAATTTG | TAGGTAAAGT | TGTAGATATC | TTATTTATCT | 7860 |
| | TTGGATTGCT | AGGTGGTGCG | GCAACATCAC | TAGCGTTAGG | TGTGCCATTA | ATTTCTGCAG | 7,920 |
| 15 | GCATAGAAAG | ATTAACTGGT | TTAGATGGTA | AAAATATGAT | TTTACGTTCG | GCCATTTTAT | 7980 |
| | TAACAATCAC | GGTTATATTT | GCCATTAGTT | CATATACAGG | ATTGAÄAAAA | GGTATTCAAA | 8040 |
| | AGTTAAGTGA | TATCAACGTT | TGGCTATCCT | TTGTACTTTT | AGCCTTTATA | TTTATTATTG | 8100 |
| 20 | GACCGACTGT | TTTTATTATG | GAAACGACAG | TGACAGGGTT | CGGAAATATG | TTGAGAGATT | 8160 |
| | TCTTTCATAT | GGCAACATGG | TTAGAACCAT | TCGGTGGTAT | TAAAGGTCGA | AAAGAAACGA | 8220 |
| | ATTTCCCACA | AGACTGGACA | ATATTCTACT | GGTCATGGTG | GTTAGTATAT | GCGCCATTTA | 8280 |
| 25 | TCGGTTTATT | TATCGCTAGA | ATTTCAAAAG | GTCGACGCCT | TAAAGAAGTC | GTGCTAGGAA | 8340 |
| | CAATTATTTA | TGGAACGCTT | GGATGCGTAT | TATTCTTTGG | TATTTTTGGT | AACTATGCTG | 8400 |
| 30 | TGTATTTACA | AATTTCTGGA | CAGTTTAATG | TAACACAATA | TTTAAATACA | CATGGTACAG | 8460 |
| 30 | AGGCAACCAT | TATTGAAGTG | GTGCATCATT | TACCATTCCC | ATCATTGATG | ATTGTACTAT | 8520 |
| | TCTTAGTATC | TGCTTTCTTA | TTCTTAGCAA | CAACATTTGA | TTCGGGTTCA | TATATTTTAG | 8580 |
| 35 | CGGCAGCATC | TCAGAAAAAA | GTGGTAGGCG | AACCATTACG | TGCCAATCGT | TTATTCTGGG | 8640 |
| | CATTTGCATT | GTGCTTATTG | CCATTTTCAT | TGATGCTAGT | TGGTGGTGAA | CGTGCATTAG | 8700 |
| | AAGTATTGAA | AACTGCTTCA | ATACTGGCAA | GTGTGCCATT | AATTGTTATT | TTTATTTTCA | 8760 |
| 40 | TGATGATATC | ATTTTTAATC | ATTTTAGGGC | GCGATAGAAT | TAAACTTGAA | ACGCGTGCTG | 8620 |
| | AAAAATTAAA | AGAAGTTGAA | CGTCGTTCAT | TGCGAATCGT | TCAAGTATCa | GAAGAAGAAC | 8880 |
| | AAGACGATAA | TTTATAATTC | AAAGCGGGTC | TGGGACGACG | AAATGAATTT | TGTGAAAATA | 8940 |
| 45 | TCATTTCTGT | TCCaTTCCCC | TTTTTTTAGT | AGCATTGTAG | GATGAACTTT | TAGGTTTTCA | 9000 |
| | TTAATGTTGT | ACTAAAAGAT | TTAATTTTTT | AGTGCTCCAA | GTACTTATTT | ATTGTATGAA | 9060 |
| 50 | GCATATTCTA | AATCGAAGTT | TGAAAGACTC | TCATTGATTA | TTAAATTAAA | TAAAGGGTAT | 9120 |
| 50 | GCGTATGTAC | AATTCAAATT | AATCGAAGGA | TGAAATAAAA | TGACTAATCA | ATTTAAAAAT | 9180 |
| | ***** | ******* | CACOTOTIACAA | macamaa aa a | 2222CTT2T2 | TOCOLACACOT | 0240 |

| | ACAGAATATT | GTTATCTATC | ATTCCGGACA | CTTAGGTGAC | TCCCAACAAG | ACATTGCATC | 9360 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | ATTAGGTGGT | GTTTCAAAAG | TATTGATGAA | TCATGATCAT | GAATCTATAG | GAGGTTCTAA | 9420 |
| 5 | TCAAGTTGAA | GCCCCTTACT | TTATACATGA | AAATGATGTG | GCTGCACTGA | AACATAAGAT | 9480 |
| | TTCTGTTCAA | AAACAATTTA | GTAATCGTGT | AATGTTGGAT | AAGGATTTAG | AAGTTATTCC | 9540 |
| 10 | CGCGCCTGGA | CATACACCAG | GGACGACACT | ATTTTTATGG | GATGATGGTC | ATCACCGTTA | 9600 |
| | CTTATTTACT | GGAGATTTTA | TATGTTTTGA | AGGGAAGAGA | TGGCGTACAG | TTATATTAGG | 9660 |
| | TTCAAGTGAT | AGAGAAAAT | CTATTCAAAG | TTTAGAGATG | GTTAAAGAAT | TAGATTTTGA | 9720 |
| 15 | TGTACTTGTA | CCTTGGGTTA | CTATCAAAGA | TGAACCGTTA | GTTTATTTTG | TAGAAAATGA | 9780 |
| | ATATGAAAAA | CGTGAACAAA | TACAAAATAT | TATTGATAGA | GTACGTGAGG | GCGAGAATAG | 9840 |
| | CTAATTGAAA | TATATTGGCG | AAgCAATGTA | ACGAATCTAA | GAAAGCCCTA | GAAAATACCT | 9900 |
| 20 | CCATAATTGA | TTGTCATATA | AAACAAAAAC | GGTAATTTCT | ATTTATTGAG | ATAGAAATTA | 9960 |
| | CCGTTTATTT | CGTGGACCTA | TTGCATTGTT | TTTATCATGC | ATAATCATCA | TTGTCGTTGT | 10020 |
| | TTGAGTCAAT | TTTAATTTTC | AGAATCAGAA | GGCTGTTCTG | GAATTGGGAA | ATATTTGAAA | 10080 |
| 25 | ATTTCACCGC | TTTCAATCGC | TTCGGTTAAC | TGTTCTAACC | ATTCGTAATA | AACATGTGTA | 10140 |
| | TGATCAAGCT | GAGCTTTAAT | TTTTTGTGCC | TCTTGTGTTT | CAGCTTCAGT | TAAATCACTG | 10200 |
| 30 | CTTTCAAGTA | ATGGATTGAT | AATAGCTTGA | GCATCTTTTA | CTGCTTCGAC | ATTGATGTCA | 10260 |
| | ATTTCACGCT | GGAATTTTT | AGTGAAAAAG | TTTCGGAAAA | AGATGAAAAA | GTCTTTCTCG | 10320 |
| | GCGATAAAAT | GTTGTTTGCG | GCTTCCTCTC | GTAAATTGTT | GTTTAACAAT | ATCAAATTCC | 10380 |
| 35 | TGCAATTTCT | TAACGCCAGC | ACTCATACTT | GGTTTGCTCA | TTTGCAATTG | ATGACGCATT | 10440 |
| | TCATCAAGCG | TCATACTGCC | TTCAAACACC | ATTGTGCCAT | ATAAGTTTCC | TACACTTCTA | 10500 |
| | TTAGTGCCAT | ACAAATCCAT | TGTCTGTCCA | ATTGAATTAA | TTACAATATC | TTTTGCTTGT | 10560 |
| 40 | TCTAATTGTT | GCTGTTTGTT | CTGAGAACGA | GTCATCATTG | CACCTCCGTA | CATCATTTTG | 10620 |
| | GTCACGTTAA | AATAAATACT | AATACATTAT | AAAACCTTTT | CTAAAAAAAG | ACATTAAAAA | 10680 |
| | TATTTAAAGC | ATTAAAGTTA | AATGTTTCGT | TAAATAAAA | TCTAACGAAC | TTACAAAACT | 10740 |
| 45 | TAATTCTTGA | GTTGTTTTGT | AAATTGACAC | ATTTTTCATT | TCTATGCTAA | CATAAGTnTG | 10800 |
| | TAAAATTcGT | TAAATAAAA | TTTAACAAAC | TTAACGGrGG | TTGTTGAAkG | Gracttttaa | 10860 |
| 50 | aACATTTATC | TCAGCGTCAA | TATATTGATG | GTGAGTGGGT | TGAAAGCGCG | AATAAAAATA | 10920 |
| 50 | CAAGAGATAT | TATCAATCCT | TACAATCAAG | AAGTGATATT | TACGGTTTCT | GAAGGGACAA | 10980 |
| | AAGAGGATGC | AGAACGTGCA | ATCTTAGCTG | CAAGACGTGC | GTTTGAGTCT | GGTGAATGGT | 11040 |

| | AACATCGCGA | AgCgTTAGCA | CGATTAGAAA | CATTAGATAC | TGGAAAAACG | TTAGAAGAAT | 11160 |
|----|------------|------------|------------|------------|------------|------------|--------|
| | CATATGCAGA | TATGGATGAT | ATTCATAATG | TGTTTATGTA | TTTTGCTGGA | TTAGCAGATA | 11220 |
| 5 | AAGACGGTGG | CGAAATGATT | GATTCACCAA | TTCCAGATAC | AGAAAGCAAA | ATTGTTAAAG | 11280 |
| | AACCAGTAGG | TGTAGTTACA | CAAATTACAC | CTTGGAATTA | TCCGTTATTA | CAAGCATCAT | 11340 |
| 10 | GGAAAATTGC | GCCAGCGCTT | GCTACGGGTT | GTTCACTAGT | TATGAAACCA | AGTGAAATTA | 11400 |
| | CACCATTAAC | AACAATACGT | GTTTTTGAAT | TAATGGAAGA | AGTTGGTTTC | CCTAAAGGAA | 114.60 |
| | CAATTAATCT | TATTCTAGGT | GCAGGTTCTG | AAGTTGGTGA | CGTAATGTCA | GGTCATAAAG | 11520 |
| 15 | AGGTTGACCT | TGTATCATTT | ACAGGTGGCA | TTGAGACTGG | TAAGCATATT | ATGAAAAATG | 11580 |
| | CTGCTAATAA | TGTTACGAAT | ATTGCCTTGG | AACTTGGCGG | TAAAAATCCA | AACATTATCT | 11640 |
| | TTGATGATGC | TGATTTTGAA | TTGGCAGTAG | ACCAAGCGTT | AAATGGTGGA | TATTTCCATG | 11700 |
| 20 | CAGGTCAAGT | TTGTTCAGCA | GGATCAAGAA | TATTAGTACA | AAACAGTATT | AAAGACAAAT | 11760 |
| | TTGAGCAAGC | ACTTATTGAT | CGCGTGAAAA | AAATCAAATT | AGGTAATGGT | TTTGATGCTG | 11820 |
| | ATACTGAAAT | GGGACCAGTG | ATTTCAACAG | AACATCGTAA | TAAGATCGAA | TCTTATATGG | 11880 |
| 25 | ATGTAGCTAA | AGCAGAAGGC | GCAACAATTG | CTGTTGGTGG | TAAACGTCCA | GATAGAGATG | 11940 |
| | ATTTAAAAGA | TGGTCTATTC | TTCGAGCCAA | CAGTCATTAC | AAATTGTGAT | ACGTCAATGC | 12000 |
| 30 | GTATTGTACA | AGAAGAGGTT | TTCGGACCTG | TCGTTACTGT | AGAAGGCTTT | GAAACTGAAC | 12060 |
| | AAGAAGCGAT | TCAATTAGCG | AATGATTCTA | TATATGGTTT | AGCAGGTGCT | GTATTTTCTA | 12120 |
| | AAGATATTGG | AAAAGCACAA | CGCGTTGCTA | ACAAGTTGAA | ACTTGGAACG | GTGTGGATTA | 12180 |
| 35 | ATGATTTCCA | TCCATATTTT | GCACAAGCGC | CATGGGGTGG | ATACAAACAA | TCAGGTATCG | 12240 |
| | GTAGAGAATT | AGGCAAAGAA | GGCTTAGAAG | AGTACCTTGT | TTCAAAACAC | ATTTTAACAA | 12300 |
| | ATAÇÃAATCC | ACAATTAGTG | AATTGGTTTA | GCAAATAAAA | ATTAGATAAG | GTGAGTGCCA | 12360 |
| 40 | TTGTAAGAAC | ACAAGACACT | CACTTTGTTT | TGTATAAGTG | GCGAAATGTT | GATTGATAAT | 12420 |
| | TTGGACTAAA | CGCAAAATGA | ATCATAGATT | ATTTCATTAC | TGTTAGTAAC | AATCGTAAAA | 12480 |
| | GGAAAAGCGA | GTGTTTTGGT | TAGCTAAGTT | TAGCAATTCA | ACGATAACCA | ATCAGCCACT | 12540 |
| 45 | AACAAATATT | TCATGCAATA | CTCACTTTGA | AATACAACAA | ACTTTGGAGG | TCATAACGAT | 12600 |
| | GAGTAACAAA | AACAAATCAT | ATGATTATGT | CATCATTGGA | GGAGGCAGTG | CAGGTTCTGT | 12660 |
| 50 | ACTAGGTAAT | CGTCTGAGTG | AAGATAAAGA | TAAAGAAGTC | TTAGTATTAG | AAGCGGGTCG | 12720 |
| 50 | CAGTGATTAT | TTTTGGGATT | TATTTATCCA | AATGCCTGCT | GCGTTAAŢGT | TCCCTTCAGG | 12780 |
| | ~~~~~~ | | | | | COCCECCEN | 12040 |

| | TCAACGTGGT AATCCAATGG ACTATGAAGG CTGGGCAGAA CCAGAAGGTA TGGAAACTTG | 12960 |
|----|---|-------|
| _ | GGATTTTGCG CACTGTTTAC CGTATTTTAA AAAATTAGAA AAAACATACG GTGCAGCGCC | 13020 |
| 5 | TTATGATAAA TTTAGAGGCC ATGATGGACC AATTAAGTTA AAACGAGGGC CAGCAACGAA | 13080 |
| | TCCTTTATTC CAGTCATTCT TTGATGCAGG TGTTGAAGCA GGCTATCATA AAACACCTGA | 13140 |
| 10 | TGTGAATGGA TTTAGACAAG AAGGTTTTGG ACCGTTCGAT AGTCAAGTAC ATCGTGGTCG | 13200 |
| | CCGAATGTCA GCTTCAAGAG CATATTTACA TCCAGCGATG AAGCGTAAAA ACTTAACCGT | 13260 |
| | TGAAACACGT GCCTTTGTAA CTGAAATTCA TTATGAAGGT AGAAGAGCAA CTGGTGTTAC | 13320 |
| 15 | GTATAAGAAA AATGGCAAAC TACATACCAT CGATGCTAAT GAAGTCATTT TGTCTGGTGG | 13380 |
| | GGCATTCAAT ACGCCACAAT TACTACAATT ATCTGGTATC GGTGATTCAG AGTTCCTAAA | 13440 |
| | ATCAAAAGGC ATTGAGCCAC GTGTTCATTT ACCTGGTGTG GGTGAAAACT TTGAAGATCA | 13500 |
| 20 | CTTAGAGG | 13508 |
| | (2) INFORMATION FOR SEQ ID NO: 121: | |
| | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 7646 base pairs | • |
| 25 | (B) TYPE: nucleic acid (C) STRANDEDNESS: double | |
| | (D) TOPOLOGY: linear | |
| 30 | | • |
| 30 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 121: | • |
| | GTAAGTATTG TCTTGATTTC CTAATAAAGT TATATCTTGT AATTCATCTT GTTGACGGCC | 60 |
| 35 | ATGTGCCATA TAAAGCGCTC CTTTAAATTT ATTTTTTAT TATTTTGGCG TCTCGGCGTG | 120 |
| | CTTTTCAAA CATGTAATAA CTTGCACCGA TAATAACGAC GTAACCTAAT GTTGCATAGA | 180 |
| | AATCTGGAGA TTCTCCGAAT AGAATAAATC CAAGTATTGC TGTGAAAATT ATAGATGCAT | 240 |
| 40 | ACGTAAAAAT AGAAATATCT TTTGCTGCTG CAAAACTATA TGCTAAAGTA ACACCAATTT | 300 |
| | GACCCACAGC GGCAGCTAAG CCAGCCCCTA ATAGATAAAG TATTTGCATC TGACTCATTG | 360 |
| | GTTCATAAGT ATATGCAGTG AAAGGTATTA AAACGATGAC AGAAAATAAG GAGAAGTAAA | 420 |
| 45 | ATACTATAGT ATATGGTGCT TYTCTTGTAC TAAGTGCTCG AACACATGTA TATGCTGATG | 480 |
| | CTGCAAAAAT ACCTGAGAAT AAGCCAGCTA ATGATGGAAT CATAGATGAT GAAAATTCAG | 540 |
| | GTTTCACTAT TAANAGCAAC CTAAAATAGC AATTATCATT GCTGTAATTT GATACTTCCT | 600 |
| 50 | TACCTTTTCA TGEAAGAAA CAATGCTTAA TAAAATCGTC CAGAAAGGAT TGAGTTTCAT | 660 |
| | TAATGAATCG GCATCACTAA GTACCATATG ATCAATCGCA TAAATATTTA ACAATACACC | 720 |

| | TGGCTGATGG | TATITATATA | TAAAAAATAA | TGGAATAAAC | ATTGCTACTA | AGTTTCGTGC | 840 |
|----|------------|------------|------------|------------|------------|------------|------|
| _ | TAATGATTTT | TGAAAAACAG | GAAGGTCACC | TGCAAGTCTG | AAAAACACTG | ACATAAAACT | 900 |
| 5 | GAAACCAATA | GCCGAAATTA | AAATGGCAAT | GATACCTTTT | ACTTTAGGAT | TCAATTTTAT | 960 |
| | CGCCTCTTTT | ATATAAAATT | AACGTATTTA | TATTAGCATA | AAACAACATG | TTGTGCATAA | 1020 |
| 10 | ATAGTTGAAA | TTTACTATAA | AAAGACTATA | ATAGACTGTA | GCGAACAAAC | GTTCTGTGTT | 1080 |
| | TATTTGTCGG | AATAATAGGG | CATTACACTT | TTATGAATGT | TTGTGTTATT | ACATAAAACA | 1140 |
| | AATATCAATT | CAGTATCAAG | CTAATAAGCT | TTTTCTTGAT | TTCTGTTGAT | ACAATTGAGA | 1200 |
| 15 | TTGACACAGA | TTTAAAAAAA | TCAAGTGATA | тстасталал | AATTTTTTTA | AATTTGTTCA | 1260 |
| | AGTTTTTCTA | ATTTAGTATT | GGTGCCTAGT | TGGAACGTTT | TACGAACATT | CGATTAGAAA | 1320 |
| | ATGGCACTTT | AAATCATAGT | GTGTCTTATG | TATAATGAAA | CACATAATAT | AGTGTTGGTG | 1380 |
| 20 | AAACGAAAAA | gACACAATAT | CTTGTGTTTT | GTATGCAAAT | GCTTTATTTA | TGAAGAAATT | 1440 |
| | ACATTTAAAA | GTAATTTAAC | ACAGAAATTT | AATAGTTATT | ATCAATTAAT | AGTCATATTT | 1500 |
| | TTAGAAAATG | TACTGAGCAA | ATGGAAGATA | TCCAATGATG | TAAACACTAC | ATATAGTGAT | 1560 |
| 25 | TTTTATACAT | TCAACCCATA | TAAGCTACTA | TTTTCTCAAA | TATAAATCTA | TGCAATTGGT | 1620 |
| | TTACATTTGA | GAAAATAAGT | AGCTTCATTA | TAGTTAATAC | AATGCTGAGA | TAACCATAGT | 1680 |
| 30 | AACCATGTTG | TTAAAGCATT | TTTTAATTGG | AATGACTACT | TTATTTAAAA | GGGTTGAAGA | 1740 |
| | AAGAAGGTGA | TCCAATGAAA | ATAATATATT | TTTCATTTAC | TGGAAATGTC | CGTCGTTTTA | 1800 |
| | TTAAGAGAAC | AGAACTTGAA | AATACGCTTG | AGATTACAGC | AGAAAATTGT | ATGGAACCAG | 1860 |
| 35 | TTCATGAACC | GTTTATTATC | GTTACTGGCA | CTATTGGATT | TGGAGAAGTA | CCAGAACCCG | 1920 |
| | TTCAATCTTT | TTTAGAAGTT | AATCATCAAT | ACATCAGAGG | TGTGGCAGCT | AGCGGTAATC | 1980 |
| | GAAATTGGGG | ACTAAATTTC | GCAAAAGCGG | GTCGCACGAT | ATCAGAAGAG | TATAATGTCC | 2040 |
| 40 | CTTTATTAAT | GAAGTTTGAG | TTACATGGAA | AAAACAAAGA | CGTTATTGAA | TTTAAGAACA | 2100 |
| | AGGTGGGTAA | TTTTAATGAA | AACCATGGAA | GAGAAAAAGT | ACAATCATAT | TGAATTAAAT | 2160 |
| | AATGAGGTCA | CTAAACGAaG | AGAAGATGGA | TTCTTTAGTT | TAGAAAAAGA | CCAAGAAGCT | 2220 |
| 45 | TTAGTAGCTT | ATTTAGAAGA | AGTAAAAGAC | AAAACAATCT | TCTTCGACAC | TGAAATCGAG | 2280 |
| | CGTTTACGTT | ATTTAGTAGA | CAACGATTTT | TATTTCAATG | TGTTTGATAT | TTATAGTGAA | 2340 |
| 50 | GCGGATCTAA | TTGAAATCAC | TGATTATGCA | AAATCAATCC | CGTTTAATTT | TGCAAGTTAT | 2400 |
| 50 | ATGTCAGCTA | GTAAATTTTT | CAAAGATTAC | GCTTTGAAAA | CAAATGATAA | AAGTCAATAC | 2460 |
| | TTAGAAGACT | ATAATCAACA | CGTTGCCATT | GTTGCTTTAT | ACCTAGCAAA | TGGTAATAAA | 2520 |

| | ACATTTTTAA | ACGCAGGCCG | TGCGCGTCGT | GGTGAGCTAG | TGTCATGTTT | CTTATTAGAA | 2640 |
|----|------------|------------|------------|------------|------------|-------------|------|
| | GTGGATGACA | GCTTAAATTC | AATTAACTTT | ATTGATTCAA | CTGCAAAACA | ATTAAGTAAA | 2700 |
| 5 | ATTGGGGGCG | GCGTTGCAAT | TAACTTATCT | AAATTGCGTG | CACGTGGTGA | AGCAATTAAA | 2760 |
| | GGAATTAAAG | GCGTAgCGAA | AGGCGTTTTA | CCTATTGCTA | AGTCACTTGA | AGGTGGCTTT | 2820 |
| 0 | AGCTATGCAG | ATCAACTTGG | TCAACGCCCT | GGTGCTGGTG | CTGTGTACTT | AAATATCTTC | 2880 |
| Ū | CATTATGATG | TAGAAGAATT | TTTAGATACT | AAAAAAGTAA | ATGCGGATGA | AGATTTACGT | 2940 |
| | TTATCTACAA | TATCAACTGG | TTTAATTGTT | CCATCTAAAT | TCTTCGATTT | AGCTAAAGAA | 3000 |
| 5 | GGTAAGGACT | TTTATATGTT | TGCACCTCAT | ACAGTTAAAG | AAGAATATGG | TGTGACATTA | 3060 |
| | GACGATATCG | ATTTAGAAAA | ATATTATGAT | GACATGGTTG | CAAACCCAAA | TGTTGAGAAA | 3120 |
| | AAGAAAAAGA | ATGCGCGTGA | AATGTTGAAT | TTAATTGCGC | AAACACAATT | ACAATCAGGT | 3180 |
| 20 | TATCCATATT | TAATGTTTAA | AGATAATGCT | AACAGAGTGC | ATCCGAATTC | AAACATTGGA | 3240 |
| | CAAATTAAAA | TGAGTAACTT | ATGTACGGAA | ATTTTCCAAC | TACAAGAAAC | TTCAATTATT | 3300 |
| | AATGACTATG | GTATTGAAGA | CGAAATTAAA | CGTGATATTT | CTTGTAAÇTT | GGGCTCATTA | 3360 |
| ?5 | AATATTGTTA | ATGTAATGGA | AAGCGGAAAA | TTCAGAGATT | CAGTTCACTC | TGGTATGGAC | 3420 |
| | GCATTAACTG | TTGTGAGTGA | TGTAGCAAAT | ATTCAAAATG | CACCAGGAGT | TAGAAAAGCT | 3480 |
| 30 | AACAGTGAAT | TACATTCAGT | TGGTCTTGGT | GTGATGAATT | TACACGGTTA | CCTAGCAAAA | 3540 |
| | AATAAAATTG | GTTATGAGTC | AGAAGAAGCA | AAAGATTTTG | CAAATATCTT | CTTTATGATG | 3600 |
| | ATGAATTTCT | ACTCAATCGA | ACGTTCAATG | GAAATCGCTA | AAGAGCGTGG | TATCAAATAT | 3660 |
| 35 | CAAGACTTTG | AAAAGTCTGA | TTATGCTAAT | GGCAAATATT | TCGAGTTCTA | TACAACTCAA | 3720 |
| | GAATTTGAAC | CTCAATTCGA | AAAAGTACGT | GAATTATTCG | ATGGTATGGC | TATTCCTACT | 3780 |
| | TCTGAGGATT | GGAAGAAACT | ACAACAAGAT | GTTGAACAAT | ATGGTTTATA | TCATGCATAT | 3840 |
| 10 | AGATTAGCAA | TTGCTCCAAC | ACAAAGTATT | TCTTATGTTC | AAAATGCAAC | AAGTTCTGTA | 3900 |
| | ATGCCAATCG | TTGACCAAAT | TGAACGTCGT | ACTTATGGTA | ATGCGGAAAC | ATTTTACCCT | 3960 |
| | ATGCCATTCT | TATCACCACA | AACAATGTGG | TACTACAAAT | CAGCATTCAA | TACTGATCAG | 4020 |
| 15 | ATGAAATTAA | TCGATTTAAT | TGCGACAATT | CAAACGCATA | TTGACCAAGG | TATCTCAACG | 4080 |
| | ATCCTTTATG | TTAATTCTGA | AATTTCTACA | CGTGAGTTAG | CAAGATTATA | TGTATATGCG | 4140 |
| | CACTATAAAG | GATTAAAATC | ACTITACTAT | ACTAGAAATA | AATTATTAAG | TGTAGAAGAA | 4200 |
| 50 | TGTACAAGTT | GTTCTATCTA | ACAATTAAAT | GTTGAAAATG | ACAAACAGCT | AATCATCTGG | 4260 |
| | TCTGAATTAG | CAGATGATTA | CACTCCTATC | тстстатттс | TCAATTATTC | ACTA ACATTA | 4320 |

| | ATGTTTTGGA | GACAAAATAT | ATCTCAAATG | TGGGTTGAAA | CAGAATTTAA | AGTATCAAAA | 4440 |
|----|----------------------|-------------|-------------|------------|---|------------|------|
| | GACATTGCAA | GTTGGAAGAC | TTTATCTGAA | GCTGAACAAG | ACACATTTAA | AAAAGCATTA | 4500 |
| 5 | GCTGGTTTAA | CAGGCTTAGA | TACACATCAA | GCAGATGATG | GCATGCCTTT | AGTTATGCTA | 4560 |
| | CATACGACTG | ACTTAAGGAA | AAAAGCAGTT | TATTCATTTA | TGGCGATGAT | GGAGCAAATA | 4620 |
| 10 | CACGCGAAAA | GCTATTCACA | TATTTTCACA | ACACTATTAC | CATCTAGTGA | AaCAAACTAC | 4680 |
| U | CTATTAGATG | AATGGGTTTT | AGAGGAACCC | CATTTAAAAT | ATAAATCTGA | TAAAATTGTT | 4740 |
| | GCTAATTATC | ACAAACTTTG | GGGTAAAGAA | GCTTCGATAT | ACGACCAATA | TATGGCCAGA | 4800 |
| 15 | GTTACGAGTG | TATTTTTAGA | AACATTCTTA | TTCTTCTCAG | GTTTCTATTA | TCCACTATAT | 4860 |
| | CTTGCTGGTC | AAGGGAAAAT | GACGACATCA | GGTGAAATCA | TTCGTAAAAT | TCTTTTAGAT | 4920 |
| | GAATCTATTC | ATGGTGTATT | TACCGGTTTA | GATGCACAGC | ATTTACGAAA | TGAACTATCT | 4980 |
| 20 | GAAAGTGAGA | AACAAAAAGC | AGATCAAGAA | ATGTATAAAT | TGCTAAATGA | CTTGTATTTA | 5040 |
| | AATGAAGAGT | CATACACAAA | AATGTTATAC | GATGATCTTG | GAATCACTGA | AGATGTGCTA | 5100 |
| | AACTATGTTA | AATATAATGG | AAACAAAGCA | CTTTCAAACT | TAGGCTTTGa | ACCTTATTTT | 5160 |
| 25 | GAGGAACGTG | AATTTAACCC | AATCATTGAG | AATGCCTTAG | ATACAACAAC | TAAAAACCAT | 5220 |
| | GACTTCTTCT | CAGTAAAAGG | TGATGGTTAT | GTATTAGCAT | TAAACGTAGA | AGCATTACAA | 5280 |
| | GATGATGACT | TTGTATTTGA | CAACAAATAA | CAATTAAATT | AAAAGACCTT | CACATGTAAA | 5340 |
| 30 | GGGAAATAGC | GATTCGTTTC | GTCTTGTCTC | CTACATGTTG | AAGGTCTTTT | TTTATGTGTA | 5400 |
| | TCTAACTCAT | TATGAGTCTG | AGTAAGAAAT | CAATGCTCTA | AGATGTACAA | TGCTATTTAT | 5460 |
| 35 | ATTGGCAGTA | GTTGGCGGGG | CCCCAACACA | GAAGCAGGCG | GAAAGTCAGC | TAACAATATT | 5520 |
| | GTGCAAGTTG | GCGGGGCCCC | AACATAGAAG | CAGGCGGAAA | GTÇAGCTAAC | AATAATGTGC | 5580 |
| | AAGTTGGCGG | GGCCCCAACA | TAAAAGCAGG | CGGAAAGTCA | GCTAACAATA | TTGTGCAAGT | 5640 |
| 40 | TCGGGCGGG | CCCCAACATA | AAGAAAAACT | TTTTCCTTTA | GAAATTATCA | CTTCCaCaTG | 5700 |
| | AGTTTTACTC | ATGTATTCCT | ATTTTTAAGT | ACACATTAGC | TGAGGCTAAT | GTTAAGAACC | 5760 |
| | ACTACTTAAT | CAATCATTAG | TAGTTTTTAT | CATTTCCACT | ATTCCCaGAC | ATCAAAATCT | 5820 |
| 45 | TAAGTGTTCT | ATTTTACTTT | AAGTAAACAA | AATACACATT | CCGAAAAATT | AAATTTCAGT | 5880 |
| | TTAATTGCAA | ATATCAATAA | AATTGACACT | AAATTATTTG | AAAGGCTATT | GAAATTATGG | 5940 |
| | TCAAAAAACG | CTACTATTAA | TGAGAAATAT | TATCAATGAT | AATGATTATC | ATTAATTTAA | 6000 |
| 50 | AGGGAGAAAA | ATTTGTAATG | AAGTATTTAT | TAAAGGGAAA | TATTTTGCTT | CTATTACTAA | 6060 |
| | ייא מייינייייני א רי | A ATTATUTOC | TENCHECATAC | CTCTCACTCA | A (************************************ | AAAGATTTAC | 6120 |

| | GTATTTTAAT | TGCTGGAAGT | TCGTTGGCTT | TAGCAGGCTT | GATAATGCAA | CAAATGATGC | 6240 |
|-----------|------------|------------|------------|------------|------------|------------|------|
| _ | AAAATAAGTT | TGTTAGTCCG | ACTACAGCTG | GAACGATGGA | ATGGGCTAAA | CTAGGTATTT | 6300 |
| 5 | TAATTGCTTT | ATTGTTCTTT | CCAACCGGTC | ATATTTTATT | AAAACTAGTA | TTTGCTGTTA | 6360 |
| | TTTGCAGTAT | TTGCGGTACG | TTTTTATTTG | TTAAAATCAT | TGATTTTATA | AAAGTGAAAG | 6420 |
| 10 | ATGTCATTTT | TGTACCGCTT | TTAGGAATTA | TGATGGGTGG | GATTGTTGCA | AGTTCACAAC | 6480 |
| | CTTCATCTCA | TTGCGCACGA | ATGCTGTTCA | AAGCATTGGT | AACTGGCTTA | ACGGGAACTT | 6540 |
| | TGCCATTATC | ACAAGTGGAC | GCTATGAAAT | TTTATATTTA | AGTATTCCTC | TTTTAGCATT | 6600 |
| 15 | GACATATCTT | TTTGCTAATC | ATTTCACGAT | TGTAGGAATG | GGTAAAGACT | TTACTAATAA | 6660 |
| | TTTAGGTTTG | AGTTACGAAA | AATTAATTAA | CATCGCATTG | TTTATTACTG | CAACTATTAC | 6720 |
| | AGCATTGGTA | GTGGTGACTG | TTGGAACATT | ACCGTTCTTA | GGACTAGTAA | TACCAAATAT | 6780 |
| 20 | TATTTCAATT | TATCGAGGTG | ATCATTTGAA | AAATGCTATC | CCTCATACGA | TGATGTTAGG | 6840 |
| | TGCCATCTTT | GTATTATTTT | CTGATATAGT | TGGCAGAATT | GTTGTTTATC | CATATGAAAT | 6900 |
| | AAATATTGGT | TTAACAATAG | GTGTATTTGG | AACAATCATT | TTCCTTATCT | TGCTTATGAA | 6960 |
| 25 | AGGTAGGAAA | AATTATGCGC | aACAATAATA | TAATAAAAA | GCTTTTAATT | GCAGTAACGT | 7020 |
| | TATTAATTAG | TATGCTGTAC | TTATTTGTAG | GTATTGATTT | TGAAATATTT | GAATATCAAT | 7080 |
| 30 | TTTCAAGTCG | TTTAAGAAAG | TTCATATTAA | TTATTTTAGT | AGGTGCTGCC | ATTGCAACTT | 7140 |
| | CAGTGGTGAT | TTTTCAAGCG | ATTACAAATA | ACCGTCTATT | GACACCATCA | ATAATGGGGT | 7200 |
| | TAGATGCAGT | TTATTTATTT | ATCAAAGTAT | TGCCAGTCTT | TTTATTTGGA | ATTCAATCGG | 7260 |
| 35 | TATGGGTTAC | TAATGTATAT | TTGAACTTTA | TATTAACACT | TATAACGATG | GTGTTATTCG | 7320 |
| | | | | | | TTTATCTTAC | 7380 |
| | TTATTGGTGT | CCTTTTAGGA | ACATTTTTTA | GAAGCATAAC | AGGTTTTATT | CAACTGATTA | 7440 |
| 40 | TGGATCCTGA | GTCATTTTTA | GCAATACAAA | GTAGTATGTT | TGCTAATTTT | AATGCTTCTA | 7500 |
| | | | | | | ATTACAATTT | 7560 |
| | TACTATTGCC | TTATTTAGAT | GTATTGCTTT | TAGGTCGTGC | TGAAGCAATT | AATCTTGGGA | 7620 |
| 45 | TATCGTATGA | AAAATTAACG | CGAATT | | | | 7646 |

(2) INFORMATION FOR SEQ ID NO: 122:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1194 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

| (xi) SEQUI | ENCE DESCRIPTION: | SEQ | ID | NO: | 122: |
|------------|-------------------|-----|----|-----|------|
|------------|-------------------|-----|----|-----|------|

| | ATGAATATAT | TTnnAAATAA | ATTATTATGG | ATTGCACCAA | TNGCCACTAT | GATTATCTTG | 60 |
|----|-------------|--------------|-------------|------------|------------|------------|------|
| 5 | GTAATCTTTT | CTTTAGCTTT | TTATCCTGCA | TATAATCCTA | AACCAAAAGA | TTTACCAATT | 120 |
| | GGTATATTAA | ACGAGGATAA | AGGTACAACG | ATTCAAGATA | AAAATGTTAA | CATTGGTAAA | 180 |
| 10 | AAATTAGAGG | ATAAATTATT | AGATAGTGAT | TCTAATAAAA | TTAAATGGGT | TAAGGTTGAT | 240 |
| | AGTGAAAAAG | ACCTTGAAAA | AGATTTGAAA | GATĆAAAAAA | TCTTTGGAGT | AGCTATTATT | 300 |
| | GATAAAGACT | TTTCAAAAGA | TGCTATGAGT | AAAACACAAA | AAGTAGTTAT | GGATAGTAAA | 360 |
| 15 | AAAGAAGAAA | TGCAACAAAA | AGTTGCTTCA | GGTGAAATTC | CGCCACAAGT | GGTTCAACAA | 420 |
| | ATGAAACAAA | AAATGGGGAA | TCAACAAGTA | GAGGTTAAGC | AGGCTAAATT | TAAAACGATT | 480 |
| | GTAAGTGAAG | GATCAAGCTT | ACAAGGTTCA | CAAATTGCAT | CAGCTGTGTT | AACTGGTATG | 540 |
| 20 | GGTGATAATA | TTAATGCTCA | AATTACGAAG | CÄAAGTTTGG | AAACATTAAC | GAGTCAAAAT | 600 |
| | GTTAAAGTCA | ATGCCGCGGA | CATCAATGGT | TTGACGAATC | CAGTAAAAGT | GGATAATGAA | 660 |
| | AAACTTAATA | AAGTTAAAGA | TCACCAAGCA | GGTGGTAATG | CACCATTCCT | AATGTTTATG | 720 |
| 25 | CCAATTTGGA | TAGGTTCAAT | CGTAACGTCT | ATCTTATTGT | TCTTTGCATT | TAGAACTAGT | 780 |
| | AACAATATCG | TCGTGCAACA | TCGTATCATT | GCtTCAATTG | GACAGATGAT | ATTTGCAGTT | 840 |
| | GTTGCAGCAT | TTGCAGGTAG | CTTTGTTTAT | ATTTATTTCA | TGCAAGGCGT | TCAAAGATTT | 900 |
| 30 | GATTTTGACC | ATCCAAATCG | TATCGCAATT | TTTGTAGCAT | TTGCGATTCT | TGGTTTCGTG | 960 |
| | GGCCTTATTT | TAGGTGTTAT | GGTATGGCTA | GGTATGAAGT | CAGTTCCAAT | TTTCTTCATT | 1020 |
| 35 | TTAATGTTCT | TTAGTATGCA | ACTTGTAACG | TTACCTAAAC | AAATGTTGCC | TGAAAGTTAT | 1080 |
| | CAAAAATATG | TATATGATTG | GAATCCATTC | ACACACTATG | CAACAAGTGT | AAGAGACTAT | 1140 |
| | TATACTTGAA | TCATCATATT | GAATTAAATA | GTACAATGTG | GATGTTTATA | GGGT | 1194 |
| 40 | (2) ÎNFORMA | ATION FOR SE | O ID NO: 12 | 23: | | | |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 558 base pairs
- (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 123:

50 GACCGACCTA TACATCCGTA TAAGTATTTC TTGATATAAG TCTTCTAAAT CATAATGATT 60 AAATCCAAAT GTTTTGATGC GTCGAATAAT TAATGGTTGT AGATCCATTA CTAACTTTTC 120

55

| GTATTTCAAA | TATTAAACTA | ACCCCTTCTA | TCTAAAATTT | AAGGTTAGTT | TAATATTGTT | 240 | | | | |
|------------|-------------------------------------|------------|------------|------------|------------|-----|--|--|--|--|
| ACATTCAAAA | TTTCAAGATG | ACGGAAATGT | CATTTCTTAT | GATGTCCTCT | TCGTATTTTT | 300 | | | | |
| TCAAATTCTG | CAAGGATTTC | AGAAGATAAC | GGAATTCGAG | TTCTTGGCTT | GTTTTCACTT | 360 | | | | |
| ATATCATCTA | ATGATTTACT | CACATCAATT | TCATTTTCTT | TTAAATCTCT | CCACATTTCG | 420 | | | | |
| CGAGATGATA | TTCTATATGC | ACCTGATCCA | AAGATAGCAT | GTTGcTCACT | CaTATCACTT | 480 | | | | |
| GTTACAACTG | TAATATGCTT | AGLATGCTTG | tCaTAAAGtT | Cataaaccat | AACGGTTCTA | 540 | | | | |
| ATGGAAACCA | ATCAGCTG | | | | | 558 | | | | |
| (2) INFORM | (2) INFORMATION FOR SEQ ID NO: 124: | | | | | | | | | |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7762 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 124:

| GCTTCAGACA | TnTGATGATA | TAATCTCTCA | TCATCGATTA | ATTCTTTTGC | AGCTTGATAC | 60 |
|------------|------------|-------------------|------------|------------|------------|-----|
| ACATATTGCT | TATTTGTTCC | AATGACTTTT | AATGTGCCAG | CTTCAACACC | TTCAGGACGT | 120 |
| TCTGTAACAC | TTCGCCAAAA | CTAAAACTGG | CTTATTAAAT | GATGGCGCTT | CTTCCTGAAT | 180 |
| TCCACCTGAA | TCTGTCAAAA | TAAAATAAGA | TTTTTTAGCA | AAATTATGGA | AATCTATACG | 240 |
| TCCAAAGGTT | CAATCAATTC | AATTCTGTCA | TGACTACCTA | AAATCTTTTG | AGCCACCTCT | 300 |
| CGAACTTTCG | GGTTTTTATG | CATTGGATAT | ACCAGTGCTA | AATCAGTATA | CTCATCTATT | 360 |
| AAGCGTCTAA | CCGCTTTAAA | TATATTTTCC | aTGGGTTTCC | CGATATTTTC | TCGTCGGTGT | 420 |
| GCTĢTCATrA | GAATGAATTT | kTtGTCATGG | TATTTATCCA | TGATGTTAGA | TTTATAATTG | 480 |
| TCATCAACTG | TATATTTCAT | AGCATCAATC | GCAGTATTAC | CAGTGACAAC | AACACTTTCT | 540 |
| GAATATTTCC | CTTCACTTAA | CAAATGCGAT | GCAGCATTTT | TAGTAGGTGC | AAAATGTAAG | 600 |
| TCAGCTAATA | CACCAACTAA | TTGTCTATTC | ACCTCTTCTG | GAAAAGGTGA | ATATTTATCA | 660 |
| TAACTTCTAA | GCCCTGCTTC | AACGTGTCCA | ATCGGCACTT | GGTTATAAAA | TGCCGCTAAA | 720 |
| CCACCTGCAA | ATGTCGTCAT | CGTATCACCA | TGTACAAGTA | CCATGTCTGG | TTTTTCTAAT | 780 |
| TGAATCACTT | GTTCTAATTG | AGTGATTGAT | TTAGAAGTTA | TCTCAGAAAG | TGTCTGTCCT | 840 |
| GATTTCATAA | TATTCAAATC | GTATTTTGGT | TTGATTTCAA | AGGTACTTAA | TACTGAATCA | 900 |
| AGCATTTCTC | TATGCTGTGC | TGTAACAACA | ACAATTGGCT | CGAGCATTTT | TTCTTGTTCC | 960 |

| | ATCTITITCA | TCAAACTACT | TATCTCCGAT | TCTTCTATTT | AGTACCAAAC | AATCTATCTC | 1080 |
|-----|------------|------------|------------|------------|------------|------------|------|
| | CAGCGTCGCC | TAACCCTGGT | GTGATATATG | CTTTGTCATT | aGCTTTTCAT | CAAGTGCAGC | 1140 |
| 5 | AATATAAATA | TCTACATCTG | GATGTGCTTC | ATGCATCTTT | TCTACGCCTT | CTGGTGCTGC | 1200 |
| • | AATTAAACAC | ATGAAGCGAA | TATTTTTAGC | GCCACGTTTC | TTCAATGAAG | TAATAGCTTC | 1260 |
| | AATTGCTGAT | GCGCCTGTTG | CTAACATAGG | ATCAACAAĆA | ATGATTTGTC | TTTCAGTAAT | 1320 |
| 10 | ATCTTGAGGT | AACTTAGCAA | AATACTCTAC | AGCCTTTAAT | GTTTCGGGAT | CTCGATATAA | 1380 |
| | ACCGATATGT | CCAACTCTGG | CTGCAGGTAC | TAAACTTAAA | ATACCATCAG | TCATACCTAA | 1440 |
| 15 | ACCAGCTCTT | aaaattggaa | CGATAGCTAA | TTTTTTACCA | GCTAATCGTT | TAGCCGTCAT | 1500 |
| , 0 | TTTAGTTACA | GGCGTTTCAA | TATCAACATC | CTGAAGCTCT | AAGTCTCTAG | TTACTTCATA | 1560 |
| | TGCCATCAAC | ATACCAACTT | CGTCTACAAG | TTCTCTAAAT | TCTTTAGTAC | CTGTATTTAC | 1620 |
| 20 | ATCTCTAATA | TAGCTTAGTT | TGTGTTGAAT | TAATGGATGA | TCGAAAACGT | GTACTTTACT | 1680 |
| | CATAAAAATT | ACTCCTATCT | TTGTGTATGT | TTATTGATAT | AGAGGATATT | CAGCTGTTAA | 1740 |
| | TTTCGCAACG | CGTTCTTTAG | CTTGTTGTAA | TTTTTCTTCA | TCTTTACTAT | TTTTCAATGC | 1800 |
| 25 | TAAACTGATG | ATTTTTGCAA | CTTCCTCAAA | AGCTTTTTCA | TCAAATCCAC | GCGTTGTTGC | 1860 |
| | AGCAGGTGTA | CCTAAACGTA | TACCACTCGT | TACAAAAGGT | TTTTCTTGAT | CGAACGGAAT | 1920 |
| | GGTATTTTTG | TTACATGTGA | TACCAACTGA | ATCTAAAGTC | TCTTCAGCTT | CTTTACCAGT | 1980 |
| 30 | AAGTCCTATA | GACCCTTTTA | CATCAACAGC | TACTAAGTGA | TTATCTGTAC | CGCCAGAAAC | 2040 |
| | AATTCTAAAT | CCTTCATTAA | TTAATGCTTC | TGCAAGAACT | TTTGCGTTTT | TAACCACTTG | 2100 |
| | TTGTTGATAC | GTTTTGAAAT | TATTTTCTAA | CGCTTCTCCA | AAAGCAACTG | CTTTtGCTgC | 2160 |
| 35 | AATAACATGC | TCAAGAGGTC | CACCTTGAAT | ACCAGGGAAA | ATTGTTTTAT | CTATGTCTTT | 2220 |
| | TTTATATTCT | TCCTTACATA | AAATCATACC | ACCACGEGGT | CCGcGTAATG | TTTTGTGTGT | 2280 |
| 40 | TGTAGTTGTT | ACAAAATCAG | CATATTCTAC | TGGATTTGGA | TGTAAACCTG | CCGCTACTAA | 2340 |
| | TCCTGCAATA | TGTGCCATGT | CTACCATTAA | CTTAGCGTTT | ACTTCATCTG | CGATTTCTTT | 2400 |
| | AAACTTTTTG | AAGTCAATTG | TTCTTGAATA | TGCTGATGCT | CCTGCCACAA | TAAGCTTAGG | 2460 |
| 45 | CTTATGCTCT | AACGCTAATT | TACGAACTTC | ATCATAATTG | ATTCGTTCTG | TGTCTTTATC | 2520 |
| | TACTCCATAT | TCAACGAAAT | TGTAGAATTT | ACCACTAAAA | TTAACAGGCG | CTCCATGTGT | 2580 |
| | CAAGTGACCA | CCATGACTCA | AATTCATACC | TAAAACTGTG | TCGCCCATTT | CTAATGCAAC | 2640 |
| 50 | TAAGTAAACA | GCCATGTTCG | CTTGTGAACC | TGAATGTGGT | TGAACATTGA | CATGTTCAGC | 2700 |
| * | тесьвьевьт | COTTTACCAC | CATCAATTCC | CATCCTTTCA | СТААСАТСТА | CAAACTCACA | 2760 |

| | TTGTGCTTCC | ATAACCGCTT | CCGATACAAA | ATTITCCGAT | GCGATTAACT | CTATGTTGCT | 2880 |
|----|------------|------------|------------|---------------|------------|--------------|------|
| | ATTTTGTCTC | TGAAATTCTC | TCTCGATTGC | TTCTGCGATA | ACTITATCTT | GCTTGGTGAT | 2940 |
| 5 | ATAAGACATA | AAATCTCCCC | TTCTTTCAAA | AAAACTTATT | GGTATTTAGC | ACGTTCGCCA | 3000 |
| | CCAATCTTTT | TCGGCCTAGA | TGTGGCAATA | GTTACAATTG | CCTGTCCTAC | TTGCTTTACT | 3060 |
| | GAGGTCCTTA | CAGGTACACA | TACATGTTTA | ATATGCATGC | CTATTAACGT | TTGACCAATA | 3120 |
| 10 | TCAATTCCAC | AAGGAACAGT | AATATGTTCG | ACCACGATCG | GATCCTTCAT | ATGCTGAAAA | 3180 |
| | GCGTATGTTG | CCAAACTCCC | TCCAGCATGT | ACATCTGGAA | CGACGGAAAC | TTCTTCCATT | 3240 |
| 15 | GTTAATGGAT | TATACTGAGA | TTTTTCTATT | GTTATCGCTC | TGTTGATATG | TTCACATCCT | 3300 |
| 13 | TGAAAAGCAA | AAGTAACGCC | TGTCTCTTTA | CTCACAACAT | CTAATGCATT | AAAAATAGTT | 3360 |
| | TCTGCAACTT | CCaTCGAACC | GACAGTCCCT | ATTTTTTCGC | CAATGACTTC | CGATGTTGAA | 3420 |
| 20 | CATCCAATTA | AACATATATC | TCCTTTATTA | AAAAAGGACA | TATCTTTTAA | TTCGTCTAAT | 3480 |
| | AACATTGTCA | AATCTTTCAT | AAAAGCCCAC | CCTTCCTAAA | AATAAAAAAG | GAATATÄGCA | 3540 |
| | AAGTGCTACA | CTCCTCTATT | ATAACTTATT | TAACTGTTAA | CATATACTAA | TTATACAGAA | 3600 |
| 25 | TTCCTACTAG | CAAATAATAT | CTTTTAATTT | TAAAATTAAA | CTTACAAGTT | CTTCATAGGT | 3660 |
| | ATGTACATAC | ATTTCTTTTG | TTCCACCGTA | TGGATCTATA | ACTTCTCCTG | CTTCTTTtAC | 3720 |
| | ATATTCATGC | AATGTGAAAA | CATGATTTTG | CAAACCAAAG | TGTGCCTCTA | TTAATTCTTT | 3780 |
| 30 | GTGCGAATAC | GACATCGTCA | AAATAATATC | TGCTTTCAAA | TCTGCTTCAG | TAAATTGTTG | 3840 |
| | CGATAAGGTC | GTTTCAGCTA | AATGATGTTC | TTCAACTAAG | TCTTCAACAT | AATTCGAAAC | 3900 |
| | ACCTTGATTG | TTCACAGCGA | ATATACCTCT | TGATTCAAAT | TGATGATTTG | GCATAACCTC | 3960 |
| 35 | TTTTGCAATA | CTTTCCGCTA | ATGGGCTACG | ACATGTGTTA | CCTGTACAAA | CGAATAAAAT | 4020 |
| | CTTCÃTAGTT | CACATCCTTT | AATAATGTGA | TTACCTGCAG | CTTTTAACAT | GCGATTCATA | 4080 |
| 40 | ATTGCTTCTG | TATTATCATT | CAGCTCAAAG | CCGTATATAT | ACGCCGCTGA | AATATTTTCA | 4140 |
| | TTTTCATCAA | GTGAATGTAA | CACATCATAA | AGATTATGAC | TTGCTTGTTT | AACATCATTG | 4200 |
| | TCATCCTGAC | ATAATTGAAT | GAATTGCGCT | TCACTTGGTA | TAAACGCCAC | CTTATTACTC | 4260 |
| 45 | GGCACAATAA | AAGCTATAGA | AGACCAATCT | TTACCGTCAT | TTCCAATTTT | GCTCTCAATA | 4320 |
| | TCTGTAATAA | TTGTAAGTGG | TGTATTGGGT | GAGTAATGCT | TATACTTCAT | ACCTGGTGCA | 4380 |
| | ATTGGCTGTT | CAGTATCATT | ATAATCAGCA | TGGGCGATAC | TATTCGGAAG | TATTTCTGTA | 4440 |
| 50 | ATCATTGCTG | CTGTTATAGA | ACCAGGTCTT | GCAATTTTAT | AAGGAAAAGA | TGTGCAATCT | 4500 |
| | | | | mamas a a a m | C11C11E1CC | 10001 m1 000 | 4560 |

| | GCACTTGGAG | CAGCTAGAGG | TTCATTTATG | ATTTGTAATA | ATTGTCTACC | TACAGAATGG | 4680 |
|----|------------|-------------|-------------|------------|------------|------------|---------|
| | CTTGGCATTC | TAACAGCAAC | TGATGATAAA | CCTCCAGAAA | CTTTTCGACA | TAGATAGCCT | 4740 |
| 5 | AGCTTTAACG | GCAATATAAA | CGAAATAGGG | CCCGGCCAGA | ATGCCTGCAT | TAACTTTTCT | 4800 |
| | ACGCGTGGAT | CCAAAGTATA | TGTAAAATCT | TTTAATTGAC | CTTTACTGTG | TATATGAACA | 4860 |
| | ATAAGCGGAT | TGTCAGATGG | ACGGCCTTTA | GCTTCATATA | TTTTAGCTAC | AGCTTCTTCA | 4920 |
| 10 | TCTGTCGCAT | TTGCTGCAAG | TCCATAAACT | GTTTCAGTTG | GTAAACCTAT | TAAACCACCG | 4980 |
| | TTTAAAACAA | TGTCTTTTAT | TTCATTAATT | TTAGGATATT | GCTGTAAATC | TTCATTATAT | 5040 |
| _ | TCTCTAACAT | CCCAAATTTT | AGTATCCAAC | TTAATCACGC | CTTTCTTATT | TATCATAATA | 5100 |
| 15 | TAAAGCAAAA | AGCTATGCAC | TTAACTAATC | ATAGCAAAGG | CATAACTTCT | AATTACCATT | 5160 |
| | TAAATGAGAC | GATTCGATCG | TGGCCATTTA | TATCTTTAAT | AATGTCGATT | TTTTTGTCAG | 5220 |
| 20 | GAAATTTATT | TAAAATTATT | GATTTAAGTG | CCTCACCTTG | ATTGTAACCA | ATTTCAAAAA | 5280 |
| | CAACTGGGCT | GCCTTTTTCC | ATAACGTGAG | GTAAATCTTC | AATGATTGAT | TCATAAATAG | 5340 |
| | CATATCCATG | GTTATCTGCA | AACAATGCCT | GATGTGGTTC | GAATCTCGTA | ACCGTTGGAG | 5400 |
| 25 | ACATCGTAAC | CATATCTTTT | TCATCTATAT | ATGGTGGATT | AGATATCAAG | CCGTTCAACT | 5460 |
| | TGATACCTTC | ATTAATTAAG | GGCTTTAATG | CATCCCCTGT | TAAAAATTGT | ATTTGTGATT | 5520 |
| | GATGCTTCTC | AGCATTATTA | CGAGCCATAT | TCATTGCTTC | AAGTGAAATA | TCAGTAGCAA | 5580 |
| 30 | TAACATTTAA | ATCCGGCTTT | TCACATTTCA | AAGTAATTGC | AAGTACACCA | CTACCCGTTC | 5640 |
| | CGATATCTAC | GATTGTTGCA | TCATCTTCTA | ACTGTTGTAA | GAAATGCAAC | ATTACTTCTT | 5700 |
| | CAGTTTCAGG | TCTTGGTATC | AAACAATTTG | AGTTTACATC | AAACGTTCTA | CCATAAAATG | 5760 |
| 35 | AGGCAAAGCC | AACTATATAC | TGTATAGGCT | CTCCTAATAA | CATACGTTGT | AATGCTAAGT | 5820 |
| | CGAACTTCAT | AATCATCGCT | TTCGGCATAT | CATCATGCAT | GTGGACTACA | AAGTCCGTAC | 5880 |
| | GCGTCCATTG | AAATACATCT | AACATTAACC | ATTCAGCTCG | TGTTTGTTCA | AACCCTTTTT | 5940 |
| 40 | GTTGTGTTAA | -ATGAATTGCT | TCATCTAACT. | TTTCTTTATA | ATTCACCATT | ATTAAGTTCT | 6.0.0.0 |
| | TTCAATTTAT | CTGTCTGCTC | TGATAAAGTC | AGTGCATCTA | TAATTTCTTC | TAAATGGCCT | 6060 |
| 45 | TCCATAATTT | GCCCTAATTT | TTGAAGCGTT | AGACCTATAC | GATGGTCTGT | TACACGGCTT | 6120 |
| | TGTGGATAAT | TATAAGTTCG | AATACGTTCT | GAACGATCAC | CAGTACCGAC | TGCTGATTTA | 6180 |
| | CGTTGTGACG | CATACTTTTG | TTGTTCTTCT | TGAACTTTCA | TATCGTATAA | ACGTGCTTTT | 6240 |
| 50 | AACACTTTCA | TTGCTTTTTC | ACGGTTTTGA | ATTTGAGACT | TCTCAGAAGA | TGTTGCAATG | 6300 |
| | ACACCAGTTG | GTAAATGGGT | AATACGTACT | GCAGAGTCAG | TTGTGTTTAC | GTGCTGACCA | 6360 |

| | ACATCTTCAA | CTTCTGGTAA | AACTGCCACT | GTAGCTGTTG | AAGTATGAAT | ACGTCCACCT | 6480 |
|----|------------|------------|------------|------------|------------|------------|------|
| | GATTCTGTTT | CAGGCACACG | TTGAACGCGG | TGCGCACCAT | TTTCAAATTT | CAATTTACTA | 6540 |
| 5 | TACGCGCCAT | TACCAGAAAC | TGAGAAACTA | ATTTCTTTGT | AACCACCATG | GTCACTTTCA | 6600 |
| | GACGCTTCTA | CTATTTCAGT | TTTGAATCCT | TGTGATTCAG | CATACTTTGA | ATACATACGC | 6660 |
| | ATTAAATCAC | CAGCAAAAAT | CGCAGCCTCA | TCACCACCTG | CTGCTGCTCT | TATTTCTACA | 6720 |
| 10 | ATAACGTCTT | TGTCATCATT | AGGATCTTTA | GGAATCAATA | ATATTTTAAG | CTCTTCTTCA | 6780 |
| | AGATTTGGAA | GTTCAGCTTT | AATACCATTA | CTCTCCTCTT | TTAACATTTC | TACTTCTTCT | 6840 |
| 15 | TTATCATCAG | TCTCACTTAA | CATTTCTTCA | ATATCAGCTA | ATTCTTCTTT | TTTAGCTTTA | 6900 |
| | TAGTTACGAT | AAACATCTAC | AGTTTTTTGT | AAATCAGCTT | GCTCTTTAGA | ATATTTACGT | 6960 |
| | AATTTATCTG | AATCATTTAC | AACATCTGGG | TCACTTAACA | GTTCATTTAA | CTGTTCGTAT | 7020 |
| 20 | CTTTCTTCTA | CAATATCTAA | TTGATCAAAC | ACTTATAATT | CCTCCTTATT | ATTATCACTA | 7080 |
| | GGTGCTACGA | TATGGTGCGC | GCGACAACGT | GGCTCATAAC | TTTCATTGGC | ACCTACTAAG | 7140 |
| | ATAATCGGAT | CATCGATTTT | AGCTGGTTTA | CCATTTATTA | ATCGTTGCGT | TCTACTAGAT | 7200 |
| 25 | GAAGAACCAC | AAACAGCACA | AACTGCTTGA | AGTTTCGTTA | CTTGTTCACT | GACAGCCATC | 7260 |
| | AATTTAGGCA | TTGGTTCGAA | CGGTTCGCCC | CTAAAATCCA | TATCTAATCC | AGCAACAATA | 7320 |
| | ACACGGTGTC | CATCTGCTGA | TAGTTTTTCT | ACTATACTTA | CAATTTCATC | GTCAAAAAAT | 7380 |
| 30 | TGCAcTTCGT | CTATTCCTAT | AACATCAACA | TTAGTTAAGT | CGTGCGTCAT | AATTTCACTT | 7440 |
| | GCTTTAGAAA | TATTAATCGC | TTCAATGGCA | TTACCATTAT | GAGAGACCAC | TTTTTCTTTA | 7500 |
| 35 | TGATATCGAT | CATCAATCGC | CGGTTTAAAT | ACAACGACTT | TTTGTTTAGC | GTATATACCC | 7560 |
| 35 | CTTCTTAGAC | GTCTTATTAG | TTCTTCGGAT | TTACCGCTAA | ACATACTACC | TGTAATACAT | 7620 |
| | TCTATCCAAC | CGGAATGGTA | AGTTTCATAC | ATTGAGAGTn | CCACCTTTTT | CAAAACATAA | 7680 |
| 40 | TCGCTTTATT | ATATCATATT | TCAAATATTC | ATAAATGTCT | TTnTCATAAT | TATATCGATA | 7740 |
| | TTGTACATGA | ACAATTATTT | TA | | | | 7762 |

(2) INFORMATION FOR SEQ ID NO: 125:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2583 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 125:

| | IAAAAAAAII | ATTATCAMIG | ATGAACTAGA | ATTGACTGAA | TTCCACCAAG | AACTTACTTA | 120 |
|-----|------------|------------|------------|------------|------------|------------|------|
| | TATTTTAGAC | AACATAnAAG | GGAATAATAA | TTATĠGTAAG | GAATTTGTTG | CAACCGTTGA | 180 |
| 5 | AGAAACATTC | GACATTGAAT | AaAGCGGGGT | GgaAGCACTA | TGAATCAATG | GGATCAGTTC | 240 |
| | TTAACACCTT | ATAAGCAAGC | GGTTGATGAG | TTGAAAGKGA | AcTTaAAGGC | ATGCGCAAAC | 300 |
| | AATATGAAGT | TGGTGAACAA | GCGTCGCCAA | TAGAATTTGT | TACTGGTCGT | GTTAAACCAA | 360 |
| 10 | TCGCTAGTAT | TATAGATAAG | GCAAACAAAC | GACAAATACC | ATTTGATAGG | TTAAGAGAAG | 420 |
| | AAATGTACGA | TATCGCTGGT | TTAAGAATGA | TGTGCCAATT | TGTTGAAGAT | ATTGATGTTG | 480 |
| . F | TCGTCAATAT | TTTAAGACAA | AGAmAAGATT | TTAAAGTAAT | TGAAGAACGA | GATTATATTC | 540 |
| 15 | GTAACACTAA | AGAAAGTGGT | TACCGCTCGT | ATCATGTCAT | TATTGAATAT | CCAATTGAAA | 600 |
| | CATTACAAGG | CCAAAAATTT | ATATTGGCTG | AGATTCAGAT | TCGTACATTA | GCAATGAATT | 660 |
| 20 | TCTGGGCAAC | GATTGAACAT | ACTTTACGAT | ATAAATATGA | TGGTGCTTAT | CCGGATGAAA | 720 |
| | TTCAACATCG | TTTGGAAAGA | GCGGCAGAAG | CAGCGTATTT | ACTTGATGAA | GAGATGTCTG | 780 |
| | AAATTAAAGA | TGAAATTCAG | GAAGCTCAAA | AATATTACAC | GCAAAAACGT | TCTAAAAAAC | 840 |
| 25 | atgaaaatga | TTAACGAGGT | GTTATAAATC | ATGCGTTATA | CAATTTTAAC | TAAAGGTGAC | 900 |
| | TCCAAGTCTA | ATGCCTTAAA | GCATAAAATG | ATGAACTATA | TGAAAGrTTT | TcGCATGaTT | 960 |
| | GaGGATrGTG | AAAaTCCTGA | AATTGTTATT | yCAGTTGGTG | GTGACGGTAC | ATTACTACAA | 1020 |
| 30 | GCATTCCATC | AGTATAGCCA | CATGTTATCA | AAAGTGGCAT | TTGTTGGAGT | TCATACAGGT | 1080 |
| | CATTTAGGAT | TTTATGCGGA | TTGGTTACCT | CATGAAGTTG | TAATTAAAA | CATCGAAATT | 1140 |
| | AATAATTCAG | AGTTTCAGGT | CATTGAATAT | CCATTGCTTG | TATTATTAAA | GAGATACAAC | 1200 |
| 35 | GACAACGGCT | ATGAAACAAG | GTATTTAGCA | TTAAATGAAG | CAACGATGAA | AACTGAAAAT | 1260 |
| | GGCŦCAACAC | TTGTTGTGGA | TGTTAACTTA | AGAGGGAAAC | ACTTTGAGCG | ATTTAGAGGC | 1320 |
| 40 | GATGGATTAT | GTGTATCAAC | ACCTTCGGGT | TCAACGGCTT | ATAACAAAGC | GCTAGGTGGC | 1380 |
| | GCACTGATAC | ATCCTTCACT | TGAAGCAATG | CAAATTACAG | AAATTGCCTC | GATAAATAAT | 1440 |
| | CGTGTGTTTA | GAACGGTAGG | ATCACCACTT | GTATTACCAA | AGCATCATAC | ATGTTTAATA | 1500 |
| 45 | TCACCAGTTA | ATCATGATAC | CATTAGAATG | ACGATAGATC | ATGTTAGTAT | CAAACATAAA | 1560 |
| | AATGTTAATT | CAATACAATA | CCGTGTAGCA | AATGAAAAAG | TGAGGTTTGC | ACGTTTTAGA | 1620 |
| | CCATTCCCAT | TCTGGAAACG | TGTGCACGAT | TCTTTCATAT | CAAGTGATGA | AGAACGATGA | 1680 |
| 50 | AATTTAAGTA | TCATATATCA | CAACAAGAAA | CTGTTAAAAC | TTTTTTAGCA | CGACATGATT | 1740 |
| | | | | | | ******* | 1000 |

| AAATACCGAG | TGTTAATTTA | ATACCTTATG | CTCGTAAGCT | AGAAGTATTG | TATGAAGATG | 1920 |
|------------|------------|------------|------------|------------|------------|------|
| CTITTATCAT | CATAGTTACT | AAACCAAACA | ATCAAAATTG | TACGCCTTCG | AGAGAACATC | 1980 |
| CTCATGAAAG | TTTAATCGAA | CAAGTACTAT | ATCATTGTCA | GGAACATGGT | GAAAATATTA | 2040 |
| ACCCACATAT | TGTTACGCGT | CTAGATCGTA | ATACAACTGG | TATTGTGATA | TTCGCTAAAT | 2100 |
| ATGGACATAT | CCATCATTTA | TTTTCTAAAG | TAAACTTGAA | AAAAATATAT | ACTTGCCTTG | 2160 |
| TATATGGTAA | AACCCATACA | TCTGGTATTA | TTGAAGCTAA | TATTAGACGG | TCAAAGGATA | 2220 |
| GGATTATAAC | TAGAGAAGTT | GCCTCGGATG | GTAAATACGC | TAAAACATCT | TATGAAGTAA | 2280 |
| TAAATCAGAA | TGATAAATAC | AGTTTATGCA | AAGTTCATTT | GCATACGGGA | CGTACACATC | 2340 |
| AAATTCGTGT | ACATTTTCAA | CATATTGGGC | ATCCAATTGT | GGGAGATTCT | TTGTATGATG | 2400 |
| GTTTTCATGA | CAAAATTCAT | GGTCAAGTAC | TGCAATGTAC | GCAAATATAT | TTTGTTCATC | 2460 |
| CAATCAATAA | GAACAATATT | TATATTACAA | TTGATTATAA | GCAATTACTT | AAATTATnCA | 2520 |
| ATCAACTCTA | ATnCACACAG | GGGGTGTAAG | TATGTCAATG | Ancacagatg | AAAAAGAGCG | 2580 |
| TGT | | | | • | · . | 2583 |
| | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 126:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1818 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 126:

| C TCAAGTTAAA AAATCnGGAT 60 | Anagatgc | GGTAAAGGAT | ACATTTAACT | ATCAAGTGAT |
|-----------------------------|----------------|------------|---------------------------|------------|
| r Attngcaatt acggatacag 120 | ACAGATTT | AAAGAGGATT | TGAATTTGTT | ATAŢĀCAATA |
| r AACATTACGT ATACCAATGT 186 | GAGTTCTT | AATGCGATTT | AGCTGGATCA | TAATAAGTAG |
| A AATTGACAAT GCAAATCATT 24 | GCGACCA | GATCAATCCC | ATTAGGTTTA | TATTAGTACC |
| A ATTAACAGCA CAAATTTTAT 30 | GAAGAACA | AAAGCGATTG | AGGATATGCT | TTGCTGATAA |
| T TATCAATAAT ATGAAATCGT 36 | ACTCGAAT | GAACAGGAAA | AAATGAAATG | TACAAGAACT |
| GATTAAAGAC GCATTGAATT -42 | gataagat | GAAGCTTTAT | TTATACGAAA | ATGAACAAAG |
| r CTTTGCTCAT CGTTTTTACA 48 | ACGTATCT | GAGTCAATGG | AATGCTTTAT | AATGGGGGGT |
| C TTGGGAAATG GATTGATAAT 54 | ATCAAGAC | GTTTTTCCAC | GAATTATCGC | TTGGTTTTTG |
| A CCACATCTAT CARCOMICCO CO | ~T-T-T A T-T-N | TTCNTCNCNC | لا له لا للمنتخلمات لا له | CAACTTTATC |

55

10

15

20

25

30

35

40

45

| | CTCATGTTAA | AGCGCCACAA | AATTGAAGCA | TTATTTTTTG | CATTAACAAT | GGCATTATCT | 720 |
|----|------------|------------|------------|------------|------------|------------|------|
| | GGAATTTTGA | ATCCAGCATT | AAAAAATATA | TTCGATAGAG | AAAGACCTAC | ATTGCTGCGT | 780 |
| 5 | TTAATTGATA | TAACAGGATT | TAGTTTTCCT | AGCGGTCATG | CTATGGGATC | AACTGCATAT | 840 |
| | TTTGGAAGTG | GTATCTATCT | ATTAAATCGA | TTAAATCAAG | GTAATTCAAA | AGGTATTCTT | 9,00 |
| | ATAGGGTTAT | GTGCAGCTAT | GATTTTATTG | ATTTCCATAT | CACGTGTATA | TCTAGGTGTA | 960 |
| 0 | CATTATCCAA | CAGATATTAT | TGCCGGCATT | ATTGGTGGAT | TATTTTGCAT | TATTTTATCA | 1020 |
| | ACGTTATTAC | TTAGAAATAA | ATTAATAAAT | TAAATAGTAA | AAAAACAAAA | GCAGTAAACC | 1080 |
| 5 | TAAAGTGTCG | TAAGGGTTTA | CTGCTTTTAT | AAAACGTTGT | TATAACGTAT | ATTGTCTTTT | 1140 |
| | ACGGGCATAT | AAnAGGGGAA | TATTTGANAA | TGACCAATCC | AACAAGAACG | AAACGTTGTG | 1200 |
| | GGGGGGATGT | TCTATGTGGT | ATTGATAATC | ATTTTCAACT | ACTATTATAC | ATTAGTGAGA | 1260 |
| 20 | ATCATTGTCA | ATTAGAAACT | AAAACTTTTT | TTGAATATTT | TTTAAGAATA | GTAAATAAAA | 1320 |
| | CGCATGATTA | CGCTATTTTA | GAAAATAAAA | AAATTTGTAT | TTCTCATTAG | AATTAGAATA | 1380 |
| | TTTAAAAGTG | ATGAGGTTTA | AACATTATAT | TGTTTACATA | CTCCTTTTGA | ATTCATACAT | 1440 |
| ?5 | TATGAAATGT | tACTTCCAAG | TTCAAAATCG | CACATTGAAA | TGATGTGTGA | AATGTTTAAA | 1500 |
| | CTACGGTCAT | tTTGTGmAAA | TAAAGrTAAT | AACTATTCAT | TTTACAATAG | TGAAAAGTCA | 1560 |
| • | GTATATGACA | ACAATTAATA | TTGCGGTAAG | GCCTTGTGTT | ACAGTATTCT | ATATTTAAGT | 1620 |
| 30 | ACTGCAATCA | GAATTAACAG | AATGCCATTA | ACTGATTATT | AAATATTTGA | GTTAATAAAT | 1680 |
| | AATTAATGAT | TGTAGCTTGA | AAAATTTAAA | ACATGGTTAT | TGATTTGTGA | TAAAATTTAA | 1740 |
| | ACGTAAACAA | ACTAATTTAA | AAAGCAACTA | TTGTATAGAA | AAATACAAAA | TTTAAAATAT | 1800 |
| 35 | ATTACCTTAT | TAGAAAA | | | | | 1818 |
| | | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 127:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12658 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

- (D) TOPOLOGY: linear

45

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 127:

| TGTTTAAACA | ATAGGGGGAA | TCTTATGATT | GAAAAATTAG | TAACCTTTTT | AAATGAGGTT | 6 |
|------------|------------|------------|------------|------------|------------|-----|
| GTTTGGAGTA | AGCCATTAGT | TTATGGTTTG | CTAATTACTG | GTGTGCTATT | TACATTGCGT | 120 |
| አጥራርር አጥጥጥ | TTCAACTTAG | асатттааа | GAAATGATTC | GATTAATGTT | TCAAGGAGAG | 18 |

55

| | GGTACAGGTA | ATATTGTCGG | TGTATCTACT | GCAATATTTA | TAGGAGGACC | TGGTGCAGTA | 300 |
|----|------------|------------|------------|------------|------------|------------|------|
| | TTTTGGATGT | GGATTACTGC | GTTTTTAGGT | GCAAGTAGTG | CTTTTATTGA | ATCTACACTT | 360 |
| 5 | GGTCAAATAT | TCAAGAGAGT | TGAAAATAAT | GAATACCGTG | GTGGACCAGC | GTATTATATT | 420 |
| | GAATATGGTA | TTGGTGGTAA | ATTTGGTAAA | ATTTACGGAA | TTATCTTTGC | TATTGTTACG | 480 |
| 10 | ATTATCTCAG | TAGGTCTATT | GCTTCCTGGT | GTGCAATCTA | ACGCTATAGC | AAGTTCTATG | 540 |
| U | CATAATGCGA | TTCATGTTCC | ACAATGGTTA | ATGGGTGGTA | TTGTTGTAGT | TATTTTGGGA | 600 |
| | TTAATTATTT | TTGGTGGTGT | ACGTATTATT | GCCAATGTTG | CAACAGCCGT | TGTACCATTT | 660 |
| 15 | ATGGCAATTA | TTTACATACT | GATGGCTGTC | ATTATCATTT | GTATCAATAT | ACAAGAAGTG | 720 |
| | CCAGCGTTAT | TTGCATTAAT | TTTCAAATCA | GCATTTGGAT | TACAATCTGC | TTTTGGTGGT | 780 |
| | ATCGTTGGCG | CAATGATAGA | GATTGGTGTT | AAACGTGGAT | TATATTCAAA | TGAGGCTGGT | 840 |
| 20 | CAAGGTACAG | GTCCACACGC | AGCAGCGGCa | gcAGaAGTAT | CACATCCAAG | TAAACAAGGT | 900 |
| | CTAGTACAAG | CATTTTCAGT | TTATATTGAT | ACATTATTTG | TATGTACTGC | AACTGCTCTG | 960 |
| | ATTATACTTA | TTTCTGGTAC | ATATAATGTG | ACTGATGGTA | CGGTTAATGC | GAATGGCACA | 1020 |
| 25 | CCGCATTTAA | TTAAAGATGG | CGGTATTTAT | GTTGAAAATG | CAACAGGTAA | AGATTATTCA | 1080 |
| | GGTACTGCGA | TGTATGCACA | AGCCGGCATt | GATAAAGCGT | TCCATGGCAG | TGGTTATCAA | 1140 |
| | TTTGATCCTA | CTTTCTCTGG | CGTAGGTTCG | TACTTTATTG | CATTTGCTTT | ATTCTTCTTT | 1200 |
| 30 | GCATTTACTA | CAATTTTGTC | GTACTACTAC | ATTACAGAAA | CAAATGTTGC | TTATTTAACG | 1260 |
| | CGTAATCAAA | ATAATCAAGT | TTCATCGATA | TTTATTAATA | TTGCTCGTGT | GATTATTTTG | 1320 |
| 25 | TTCGCTACAT | TTTACGGTGC | AGTTAAAACA | GCTGATGTAG | CATGGGCATT | CGGTGATTTA | 1380 |
| 35 | GGTGTAGGTC | TAATGGCTTG | GTTAAATATC | ATTGCGATTT | GGATTTTACA | TAAGCCTGCC | 1440 |
| | GTAAATGCTT | TAAAAGATTA | TGAAATTCAA | AAGAAACGTT | TAGGCAACGG | TTATAATGCA | 1500 |
| 40 | GTTTATCAAC | CTGATCCGAA | TAAATTACCT | AATGCTGTCT | TTTGGTTGAA | GACGTATCCA | 1560 |
| | GAACGTTTAA | AACAAGCACG | TGCCAAAAAG | TAATCTACTT | TTGTTTATAG | TATATGTAGT | 1620 |
| | GATCATTIGA | TAAAAAAGAA | AAGTATTGAG | AATTTTAGGt | GCTCAGAAAT | TTGAATTTTA | 1680 |
| 45 | AAAATATAGT | GTCTCTTGGT | ACAATAACAA | TACAACTACT | AGGGGCACTT | TTTTATGTCA | 1740 |
| | GAATTTAAAA | CTGGTAAGAT | TAATAAACAT | GTTTTATATA | GTAATATTTT | AAATAGAGAT | 1800 |
| | GTCACGTTAA | GTATTTATTT | ACCAGAATCT | TATAATCAAC | TTGTTAAATA | TAATGTCATT | 1860 |
| 50 | CTTTGCTTTG | ACGGATTAGA | TTTTTTACGT | TTCGGGAGAA | TACAACGTAC | ATATGAATCG | 1920 |
| | TTAATCAAAG | AAGCGCGTAT | TGATGATGCG | ATCATTGTTG | GATTCCATTA | TGAAGACGTT | 1980 |

| | GTCGGTAAAG | AAATATTGCC | ATTTATTGAC | TCGACGTTTT | CTACACTGAA | AGTAGGTAAT | 2100 |
|----|-------------|------------|------------|------------|------------|-------------|------|
| | GCAAGGTTAT | TAGTAGGGGA | TAGTTTAGCG | GGTAGTATTG | CCTTATTAAC | GGCGTTGACC | 2160 |
| 5 | TATCCAACGA | TTTTTAGTCG | TGTAGCAATG | TTAAGTCCAC | ATTCAGATGA | AAAAGTATTA | 2220 |
| | GATAAGCTAA | ATCAATGTGC | AAATAAAGAA | CAATTGACAA | TTTGGCATGT | CATTGGTCTA | 2280 |
| | GATGAAAAAG | ATTTTACTTT | ACCAACAAAT | GGTAAGCGTG | CCGATTTCTT | AACACCGAAT | 2340 |
| 10 | AGAGAATTAG | CTGAACAAAT | TAAGAAATAT | AATATAACTT | ATTATTACGA | TGAATTTGAT | 2400 |
| | GGTGGTCACC | AATGGAAAGA | TTGGAAACCA | TTGCTGTCAG | ATATATTATT | GTATTTTTTA | 2460 |
| | AGTAAAAACA | CAGATGATCA | ACTTTATGAA | TAATTTACAT | TAGTAGATTT | AGTATGAATT | 2520 |
| 15 | GTCTTCATAT | AGTCTGGTCT | ATAATATAAT | TTATAAAAGA | TTTTACTGTT | TAATTTAATT | 2580 |
| | TAAATTTGAC | GAAATTGCAA | AAGATGTATA | ATGAATTATT | TTTAATGTAA | CGGTTTTCAA | 2640 |
| 20 | AGAAATTTGA | TATAATAGCA | ATAGGTTAAA | CAAAGGAGGA | ATTCAGATGA | TTTTAGGATT | 2700 |
| | AGCATTAATT | CCATCAAAGT | CATTTCAAGA | AGCGGTGGAT | TCTTACCGTA | AAAGATATGA | 2760 |
| | TAAACAGTAT | TCACGAATTA | AACCACATGT | GACAATTAAA | GCGCCATTTG | AAATTAAAGA | 2820 |
| 25 | TGGTGATTTA | GATTCTGTCA | TTGAACAGGT | TAGAGCTCGT | ATTAATGGTA | TACCAGCAGT | 2880 |
| | AGAAGTTCAT | GCTACAAAAG | CTTCTAGCTT | CAAACCAACG | AACAATGTGA | TTTACTTTAA | 2940 |
| | AGTTGCGAAG | ACGGACGACT | TAGAAGAATT | GTTTAATCGC | TTTAATGGAG | AAGATTTCTA | 3000 |
| 30 | TGGAGAAGCT | GAACATGTTT | TTGTGCCACA | CTTTACAATA | GCACAAGGAC | TATCTAGCCA | 3060 |
| | AGAATTCGAA | GATATTTTTG | GTCaAGTAGC | ATTAGCTGG | GTAGACCATA | AAGAAATTAT | 3120 |
| | CGATGAATTA | ACTTTGTTAC | GTTTTGACGA | TGACGAAGAT | AAATGGAAAG | TTATTGAAAC | 3180 |
| 35 | GTTTAAATTA | GCTTAAGTAA | CATAATAGTA | TTGTTAATCG | TAGTATGTTT | GAATTAATAA | 3240 |
| | GAAAATGGTC | ATTTTTATTG | AATGTAATAA | AAATGACCAT | TTTCTTTATT | TTAAAATACG | 3300 |
| | TTTTAACCTT | ACTTAGCTTT | TTCTCTATTT | ACTATAAAGT | rGCTTCCATA | AAATACAGCT | 3360 |
| 40 | -AAGACTAAAA | AGATTAATGC | CGAGAAATAA | AATGTATTGT | TTAAATTGTT | -GGTAAATTGT | 3420 |
| | GTAATTAATC | CGCCAAATAA | TGGCCCTATC | ATTGAGCCGA | ATCCTTGGAT | ACTATTAAAA | 3480 |
| 45 | ACACCCCAAG | TTTCTTCTTG | TTCATCTGAT | TTGATAAATC | GTGCCATAAA | GGTATTCCAT | 3540 |
| | GCTGGTAATA | AGATGCCATA | CATTAGACCG | ATAGCTAAAG | CGATAATCCA | CAAGATGTGA | 3600 |
| | ATATTAACAA | TCATAGATAG | AGTAAAAATT | AATATCATGT | ATAAAATAAA | TCCGCTTAGA | 3660 |
| 50 | ATAACACCAT | ACATAAAGTT | TCTGCTGCGG | TTATCTATTA | GTTTCGATAA | AAATAGCATC | 3720 |
| | GAAACTGCAC | AGCCGATACC | ACCAATAATG | ATTGCAACAG | TATATTCAAT | TGTGCTTACG | 3780 |

| | TGTAAAAGAA | TACCAGGGAA | CAACAATAAA | TGGcGCTTTG | TCACATCAAC | AATTIGTCTC | 3900 |
|----|------------|------------|---------------|------------|------------|------------|------|
| | AATTGAGCTT | TAACTGGACG | AGTATTATAA | TTTGTTAACT | TTACATCGAC | AAAATAATAT | 3960 |
| 5 | AATATCCATG | CAATTAAAAC | GACTAAAGAC | ATCATGAAGG | CAAAGCGTGT | TGGGTGCACT | 4020 |
| | TTGATAAGTA | GATTCATAAA | AACCATACCT | ACCAATAGGC | CTAACAACCA | TGAAAAATAA | 4080 |
| | ACATAGCCCA | TTTGTTTGCC | ACGTTTATCT | TCTTCAACAC | TGGATAACAT | AATGACCCAA | 4140 |
| 10 | ATAGGACTAA | CTGCAATACC | GAGCATCATA | GCACTAAATA | TGATTACAAA | AGGTGATGCT | 4200 |
| | GGAAACCAAA | таастааааа | TAAACTIGTA | AATGCTAAAA | TAAATCCAGT | CGTTAAAACG | 4260 |
| 15 | ATTTTTGTGC | CGAATTTTTT | CAGTAAAAAT | CCTATAACAA | AGTTTGTAGA | TGCATCAGCA | 4320 |
| | ATAAAATGTA | TTGAAAATGC | TAGAGACGTT | ATTGCTACAG | CAATGGATGT | AACTGTTGGC | 4380 |
| | AAGAAATTAA | TATAGCTTAG | GATATACATG | CCTCTCGCAA | ATTCCATTAA | AAATAAGATA | 4440 |
| 20 | ATAAGCaTTA | AAATGAAATT | TTTATGATTA | GCGTAATTAT | TTAACGAAGA | ATCTTGCATA | 4500 |
| | TAAAGGAACC | TTTCCATAAA | TCTCTTGTGG | TTGTGATGAA | TGACCGATTA | AATCAAGTAA | 4560 |
| | GTCTCGACAT | ATTGTCTGTG | TAGCATACTT | AATTTTATCT | TGTTCCATTG | TACTAATCAT | 4620 |
| ?5 | GTTAGTTAAT | TGCTCATTAC | CGTTAGTTAA | ACTTGCTACA | ATTTTTATTG | CTTCTTCTGG | 4680 |
| | AGTATCAGCG | ATTTTACCAA | AACCTTTTTC | TTCAAAGTAA | AGGGCATTTT | CAAGCTCTTG | 4740 |
| | ACCAGGTGCA | GGATTTAGGA | AAATCATTGG | AATACAACGG | GCGAAACCTT | CAGTTATTGT | 4800 |
| 30 | GATACCACCA | GGTTTCGTAA | TCATAAGTTG | ACTTGATGCC | ATCCATTCAT | TCATGTGTTT | 4860 |
| | GGTATAACCT | AGAATCAATA | CATTCTCGTT | AGATTTAAAC | TTAGCTGTTA | AAGAACGCTT | 4920 |
| | TAGCTCTTTG | CTCTTACCAC | AAATCATAAC | TACTTGTGCA | TTTGCaCTTT | tCGCTAATAT | 4980 |
| 35 | ATCAGTAATC | ATCGTGTCAA | AACCTTTAGA | TACACCAAAT | GCACCAGCTG | aCATTAAAAT | 5040 |
| | AGTITGCTTA | TCTGGATCTA | AGTTGTTGTC | TATTAACCAC | TGCTTTTGAT | TAATAGGCGT | 5100 |
| 10 | TTCAAATTTG | TTATCAATAG | GAATACCTGT | CaCTTTAACT | GTTGAAGGAT | CAATACCTAC | 5160 |
| | GTCTATGAAG | TCTTGTTTCG | TTTCTTTTGT | TGCCACATAA | TATCTTGTTG | AATACGGCGT | 5220 |
| | AATCCAGTTT | TTATGTAAGC | GATAGTCTGT | CATCACTGTA | GCAACTGGAA | TATTAATGTT | 5280 |
| 15 | AAATTGCTCA | GTTAGTACCG | ACATAACTGG | TGTAGGAAAC | GTTAATAATA | TTAAATCTGG | 5340 |
| | CTTTTCTTTT | ATCAATAAAT | TAATTAACTT | ATTAAGTCCA | TAGTATTTGT | AAAAACATTT | 5400 |
| | GTCTAGTTTA | TCTGGGCGGC | TGTAATAAAA | CCCTTTGTAC | ATATTTCTAA | AAATTTTAAA | 5460 |
| 50 | GCTATTGATA | TACCATTTTT | TACAAATAGA | AGTCAAAATT | GGATGAGCTT | CCATAAATAA | 5520 |
| | ATCTTCTCA | ATCACCCTTA | а атусстста с | ATTCATATCA | TTAAGTTGAT | TAACGATACT | 5580 |

| | TTGAGTAACC | ATTAATAGCC | ACCCTCCGTT | AGTTTGAAAA | TTTTATTTAA | GTGTAACTTA | 5700 |
|----|--------------|------------|------------|------------|------------|------------|-------|
| | TTTTACGGCA | TTATAAAAGA | AATAAAGACG | CAAAGTCGTT | ACATTTATAG | CAATTTTAAT | 5760 |
| 5 | CTATAGATGA | ATTGATACAA | AATAAAACGT | TATTTTATAA | AGCAATTTAT | TGTTCTATGT | 5820 |
| | TTTATTTGTA | TATTTAAAAT | TATCCAGTAT | ACAATTATAG | CATATTTTTG | GAAACAATTA | 5880 |
| | TGATATTATA | CCATGTTACA | AGATGGTTTT | AATAATTTAA | GATGAGCCAT | AATTGTAAAA | 5940 |
| 10 | CTAATTCATA | ATACCGTATG | TTTTATTTTT | AATAGTAGAA | ATTAGAAAAT | GCTGATTAGT | 6000 |
| | AGGATATAAC | AGTGAAATTA | TAAATTTATT | AACATCAACA | AAACGTGTAT | AATAAACATA | 6060 |
| 15 | TTGTAGAAAA | AGGAGCGGTT | CAGTTTGGAT | GCAAGTACGT | TGTTTAAGAA | AGTAAAAGTA | 6120 |
| | AAGCGTGTAT | TGGGTTCTTT | AGAACAACAA | ATAGATGATA | TCACTACTGA | TTCACGTACA | 6180 |
| | GCGAGAGAAG | GTAGCATTTT | TGTCGCTTCA | GTTGGATATA | CTGTAGACAG | TCATAAGTTC | 6240 |
| 20 | TGTCAAAATG | TAGCTGATCA | AGGGTGTAAG | TTGGTAGTGG | TCAATAAAGA | ACAATCATTA | 6300 |
| | CCAGCTAACG | TAACACAAGT | GGTTGTGCCG | GACACATTAA | GAGTAGCTAG | TATTCTAGCA | 6360 |
| • | CACACATTAT | ATGATTATCC | GAGTCATCAG | TTAGTGACAT | TTGGTGTAaC | GGGTACAAAT | 6420 |
| 25 | GGTAAAACTT | CTATTGCGAC | GATGATTCAT | TTAATTCAAA | GAAAGTTACA | AAAAAATAGT | 6480 |
| | GCATATTTAG | GAACTAATGG | TTTCCAAATT | AATGAAACAA | AGACAAAAGG | TGCAAATACG | 6540 |
| | ACACCAGAAA | CAGTTTCTTT | AACTAAGAAA | ATTAAAGAAG | CAGTTGATGC | AGGCGCTGAA | 6600 |
| 30 | TCTATGACAT | TAGAAGTATC | AAGCCATGGC | TTAGTATTAG | GACGACTGCG | AGGCGTTGAA | 6660 |
| | TTTGACGTTG | CAATATTTTC | AAATTTAACA | CAAGACCATT | TAGATTTTCA | TGGCACAATG | 6720 |
| | GAAGCATACG | GACACGCGAA | GTCTTTATTG | TTTAGTCAAT | TAGGTGAAGA | TTTGTCGAAA | 6780 |
| 35 | GAAAAGTATG | TCGTGTTAAA | CAATGACGAT | TCATTTTCTG | agtatttaag | AACAGTGACG | 6840 |
| | CCTȚATGAAG | TATTTAGTTA | TGGAATTGAT | GAGGAAGCCC | AATTTATGGC | TTAAAAATT | 6900 |
| | CAAGAATCTT | TACAAGGTGT | CAGCTTTGAT | TTTGTAACGC | CTTTTGGAAC | TTACCCAGTA | 6960 |
| 40 | -AAATCGCCTT- | ATGTTGGTAA | GTTTAATATT | TCTAATATTA | TGGCGGCAAT | GATTGCGGTG | -7020 |
| | TGGAGTAAAG | GTACATCTTT | AGAAACGATT | ATTAAAGCTG | TTGAAAATTT | AGAACCTGTT | 7080 |
| 45 | GAAGGCCGAT | TAGAAGTTTT | AGATCCTTCG | TTACCTATTG | ATTTAATTAT | CGATTATGCA | 7140 |
| | CATACAGCTG | ATGGTATGAA | CAAATTAATC | GATGCAGTAC | AGCCTTTTGT | AAAGCAAAAG | 7200 |
| | TTGATATTTT | TAGTTGGTAT | GGCAGGCGAA | CGTGATTTAA | CTAAAACGCC | TGAAATGGGG | 7260 |
| 50 | CGAGTTGCCT | GTCGTGCAGA | TTATGTCATT | TTCACACCGG | ATAATCCGGC | AAATGATGAC | 7320 |
| | CCGAAAATGT | TAACGGCAGA | ATTAGCCAAA | GGTGCAACAC | атсавааста | TATTGAATTT | 7380 |

| | GTTTTAGCAT | CAAAAGGAAG | AGAACCATAT | CAAATCATGC | CAGGGCATAT | TAAGGTGCCA | 7500 |
|----|-------------|------------------|------------|-----------------|-----------------------|---------------|------|
| | CATCGAGATG | ATTTAATTGG | CCTTGAAGCA | GCTTACAAAA | AGTTCGGTGG | TGGCCCTGTT | 7560 |
| 5 | GATTAATAAA | AGATTTATTG | ATGAAGGTAA | AACTATTGAT | GTTTATTTAT | TCGAAGCATT | 7620 |
| | AAATAACCAG | ATAATCATTG | CTATACCAGA | TTGGTTTTGG | TCATATCAGA | TGGCAATGAC | 7680 |
| | ATTAGATGAA | GAAACTTGTT | TTGAAGCAAT | ACTCATGCAA | TIGTTIGTTT | TTAAAGAAGA | 7740 |
| 10 | GGAAGAGGCA | GAATCGATTG | CATCACAACT | AACAGATTGG | ATAGAAACAT | ATAAAAAGGA | 7800 |
| | GAAAGACTAA | TGAACTTAAA | GCAAGAAGTT | GAGTCTAGAA | AGACTTTTGC | GATTATTTCA | 7860 |
| 15 | CATCCCGATG | CAGGGAAAAC | AACGTTAACT | GAAAAACTAT | TGTACTTCAG | TGGTGCTATT | 7920 |
| | CGTGAAGCGG | GTACAGTTAA | AGGGAAGAAG | ACTGGTAAAT | TTGCGACAAG | TGACTGGATG | 7980 |
| | AAAGTTGAAC | AAGAGCGTGG | TATTTCTGTA | ACTAGTTCAG | TAATGCAATT | TGATTACGAT | 8040 |
| ?0 | GATTATAAAA | TCAATATCTT | AGATACACCA | GGACATGAAG | ACTTTTCAGA | AGATACGTAT | 8100 |
| | AGAACATTAA | TGGCAGTTGA | CAGTGCTGTC | ATGGTCATAG | ACTGTGCAAA | AGGTATTGAA | 8160 |
| | CCACAAACAT | TGAAGTTATT | TAAAGTTTGT | AAAATGCGTG | GTATTCCAAT | CTTTACATTC | 8220 |
| ?5 | ATTAATAAAT | TAGACCGAGT | AGGTAAAGAA | CCATTTGAAT | TATTAGATGA | AATCGAAGAG | 8280 |
| | ACATTAAATA | TTGAAACATA | CCCTATGAAT | TGGCCAATTG | GTATGGGACA | AAGTTTCTTT | 8340 |
| | GGCATCATTG | ATAGAAAGTC | TAAAACAATT | GAACCATTTA | GAGATGAAGA | AAATATATTA | 8400 |
| BO | CATTTGAATG | ATGATTTTGA | GTTGGAAGAA | GATCATGCAA | TTACAAATGA | TAGTGATTTT | 8460 |
| | GAACAAGCGA | TTGAAGAATT | AATGTTGGTT | GAAGAAGCGG | GTGAAGCCTT | TGATAATGAC | 8520 |
| | GCGCTGTTGA | GTGGAGACTT | AACACCTGTA | TTTTTCGGTT | CAGCTTTAGC | TAACTTTGGT | 8580 |
| 35 | GTACAAAATT | TCTTAAATGC | ATATGTTGAT | TTTGCGCCAA | TGCCAAATGC | GAGACAAACA | 8640 |
| | AAAGAAGACG | TTGAAGTAAG | CCCGTTTGAT | GATTCATTTT | CAGGATTTAT | CTTTAAAATT | 8700 |
| | CAAGCCAACA | TGGACCCTAA | ACACCGTGAT | AGAATTGCCT | TTATGCGTGT | CGTTAGTGGT | 8760 |
| 10 | GCATTTGAAC | GTGGTATGGA | TGTTACTTTG | CAACGTACTA | ATAAAAAGCA | AAAGATCACA | 8820 |
| | CGTTCAACGT | CATTTATGGC | AGACGATAAA | GAAACTGTGA | ATCATGCTGT | AGCAGGCGAT | 8880 |
| 15 | ATCATTGGAC | TATATGATAC | TGGTAATTAT | CAAATTGGAG | ATACTTTAGT | TGGTGGAAAA | 8940 |
| | CAAACCTACA | GTTTCCAAGA | TTTACCACAA | TTTACGCCAG | AAATTTTTAT | GAAAGTTTCT | 9000 |
| | GCTAAAAAACG | TCATGAAACA | GAAGCATTTC | CATAAAGGTA | TTGAACAATT | AGTACAAGAA | 9060 |
| 50 | GGTGCGATTC | AATACTATAA | AACATTACAC | ACAAACCAAA | TTATTTTAGG | TGCTGTTGGT | 9120 |
| | CAGTTACAAT | יייייייט ע פאריי | CCARCATACA | ATCA A B A A CC | 3 3 T 3 T 3 T 3 T 7 T | TO A DOCUMENT | 0100 |

| | AAGATGAACA | CATCAAGATC | GATTTTAGTG | AAAGATAGAT | ATGACGATTT | AGTATTCTTA | 9300 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | TTTGAAAATG | AATTTGCAAC | AAGATGGTTT | GAAGAGAAAT | TCCCTGAAAT | TAAATTGTAT | 9360 |
| 5 | AGTTTACTTT | AACAGCTCAA | TTGTATAATC | GAATTTGTTA | CATTAAAAAT | AATTGTTTCG | 9420 |
| | TTGAAGAAAA | ATAAATTGTA | TATTTTAAAA | GAAAAAGGTA | TACTATGATG | TATCAAATGA | 9480 |
| | ATAACCTATG | GCATTTTGTC | AGAGGGGAGT | AACTTAAGAA | TCATGACCGT | ATAAATGaTT | 9540 |
| 10 | CGACACTTTA | TCGTCATTAC | GArGATATCT | TCCGGTAAAG | TGGGCAATTT | AAATTGCTTA | 9600 |
| | GTGAGACCTT | TGCTATTTAT | TTAGCATAGG | TCTTTTTGTT | TGTACTTAAC | TTATTTATTT | 9660 |
| 15 | AAAGGAGTTG | TACATGTTAA | TGGATCCAAG | TTTGATCTTA | CCTTATTTAT | GGGTACTTGT | 9720 |
| | CGTTTTAGTA | TTTTTAGAAG | GCTTATTAGC | AGCAGATAAC | GCGATTGTTA | TGGCTGTAAT | 9780 |
| | GGTTAAGCAC | TTACCACCCG | AACAACGTAA | AAAAGCTTTG | TTTTACGGTT | TGTTAGGTGC | 9840 |
| 20 | ATTTGTATTT | AGATTTTTAG | CATTATTCTT | AATTAGTATT | ATCGCGAACT | TTTGGTTTAT | 9900 |
| | TCAAGCTGCA | GGAGCGGTTT | ACTTAATTTA | TATGTCAATC | AAAAATCTGT | GGCAGTTCTT | 9960 |
| | TAAACACCCA | GAAATTGAAA | GTCCTGAAGC | TGGAGATGAT | CATCATTATG | ATGAATCTGG | 10020 |
| 25 | TGAAGAGATT | AAAGCAAGTA | ACAAATCATT | CTGGGGAACT | GTGTTGAAAA | TAGAATTTGC | 10080 |
| | AGATATCGCA | TTTGCCATTG | ATTCTATGCT | TGCTGCTTTA | gCTATTGCTG | TAACACTTCC | 10140 |
| 30 | TAAAGTTGGT | ATTCACTTTG | GTGGTATGGA | CTTAGGTCAG | TTCGTAGTCA | TGTTCCTAGG | 10200 |
| 00 | TGGAATGATT | GGTGTTATTC | TAATGCGTTA | TGCAGCAACA | TGGTTTGTAG | AGCTATTAAA | 10260 |
| | CAAATATCCA | GGACTTGAAG | GTGCAGCCTt | CGCGATCGTT | GGTTGGGTAG | GTGTTAAATT | 10320 |
| 35 | AGTTGTCATG | GTATTAGCGC | ACCCAGACAT | CGCTGTATTG | CCTGAGCACT | TCCCACATGG | 10380 |
| | CGTATTATGG | CAATCTATTT | TCTGGACAGT | ACTAATTGGA | TTÁGTAATTA | TCGGTTGGTT | 10440 |
| | AGGTTCAGTT | GTTAAAAATA | AAAAATCGCA | TAAATAATTG | ATGTGAAGCG | GACAATCTTA | 10500 |
| 40 | ATTTAGTTTA | AGGTTGTCCT | TTTTCATTTA | ATTGAGTGAT | TTATGAAAAA | TGGATTTTGA | 10560 |
| | AGAATGTGAA | TCAAAAGATG | CGATATAGTA | TTAAGAAAAT | GTGCCTTTTA | TATTTAGCAT | 10620 |
| | TTTTTCAATA | GAAATTATAT | AGATTTTAAA | GCAAATTAGG | TGTTAATGTG | TCATAATGAT | 10680 |
| 45 | AAGTGATTTT | ATTGAATGGA | GTGGACATTA | GTGGATATTG | GTAAAAAACA | TGTAATTCCT | 10740 |
| | AAAAGTCAGT | nACCsaCGTA | AGCGTCGTGA | ATTCTTCCAC | AACGAAGACA | GAGAAGAAAA | 10800 |
| 50 | TTTAAATCAA | CATCAAGATA | AACAAAATAT | AGATAATACA | ACATCAAAAA | AAGCAGATAA | 10860 |
| | GCAAATACAT | AAAGATTCAA | TTGATAAGCA | CGAACGTTTT | AAAAATAGTT | TATCATCGCA | 10920 |
| | TTTAGAACAG | AGAAACCGTG | ATGTTAATGA | GAATAAAGCT | GAAGAAAGTA | AAAGTAATCA | 10980 |

| | MATICATIA | GATICAGIGG | ACCAAGATAC | AGAGAAATCA | AAATATTATG | AGCAAAATTC | 11100 |
|----|--------------|------------|------------|------------|------------|------------|-------|
| | TGAAGCGACT | TTATCAACTA | AATCAACCGA | TAAAGTAGAA | TCAACTGAAA | TGAGAAAGCT | 11160 |
| 5 | AAGTTCAGAT | AAAAACAAAG | TTGGTCATGA | AGAGCAACAT | GTACTTTCTA | AACCTTCAGA | 11220 |
| | ACATGATAAA | GAGACTAGAA | TTGATTCTGA | GTCTTCAAGA | ACTGATTCAG | ACAGCTCGAT | 11280 |
| 0 | GCAGACAGAG | AAAATAAAA | AAGACAGTTC | AGATGGAAAT | AAAAGTAGTA | ATCTGAAATC | 11340 |
| | TGAAGTAATA | TCAGACAAAT | CAAATACAGT | ACCAAAATTG | TCGGAATCTG | ATGATGAAGT | 11400 |
| | AAATAATCAG | AAGCCATTAA | CTTTACCGGA | AGAACAGAAA | TTGAAAAGAC | AGCAAAGTCA | 11460 |
| 5 | AAATGAGCAA | ACAAAAACCT | ATACATATGG | TGATAGCGAA | CAAAATGACA | AGTCTAATCA | 11520 |
| | TGAAAATGAT | TTAAGTCATC | ATATACCATC | GATAAGTGAT | GATAAAGATA | ACGTCATGAG | 11580 |
| | AGAAAATCAT | ATTGTTGACG | ATAATCCTGA | TAATGATATC | AATACACCAT | CATTATCAAA | 11640 |
| 20 | AACAGATGAC | GATCGAAAAC | TTGATGAAAA | AATTCATGTT | GAAGATAAAC | АТАААСАААА | 11700 |
| | TGCAGACTCG | TCTGAAACGG | TGGGATATCA | AAGTCAGTCA | ACTGCATCTC | ATCGTAGCAC | 11760 |
| | TGAAAAAAGA | AATATTTCTA | TTAATGACCA | TGATAAATTA | AACGGTCAAA | AAACAAATAC | 1182 |
| 25 | AAAGACATCG | GCAAATAATA | ATCAAAAAA | GGCTACATCA | AAATTGAACA | AAGGCCCCC | 1188 |
| | TACGAATAAT | AATTATAGTG | ACATTTTGAA | AAAGTTTTGG | ATGATGTATT | GGCCTAAATT | 11940 |
| 10 | AGTTATTCTA | ATGGGTATTA | TTATTCTAAT | TGTTATTTTG | AATGCCATTT | TTAATAATGT | 12000 |
| | GAACAAAAAT | GATCGCATGA | ATGATAATAA | TGATGCAGAT | GCTCaAAAAT | ATACGACAAC | 12060 |
| | GATGAAAAAT | GCCAATAACA | CAGTTAAATC | GGTCGTTACA | GTTGAAAATG | AAACATCAAA | 12120 |
| 15 | AGATTCmTCA | TTACCTAAAG | ATAAAGCATC | TCaAGACGAA | GTGGGATCAG | GTGTTGTATA | 12180 |
| | TAAAAAATCT | GGAGATACGT | TATATATTGT | TACGAATGCA | CACGTTGTCG | GTGATAAAGA | 12240 |
| | AAATÉaAAAA | ATAACTTTCT | CGAATAATAA | AAGTGTTGTT | GGGAAAGTGC | TTGGTAAAGA | 12300 |
| 10 | TAAATGGTCA | GATTTAGCTG | TTGTTAAAGC | AACTTCTTCA | GACAGTTCAG | TGAAAGAGAT | 12360 |
| | AGCTATTGGA | GATTCAAATA | ATTTAGTGTT | AGGAGAGCCA | ATATTAGTCG | TAGGTAATCC | 12420 |
| | ACTTGGTGTA | GACTTTAAAG | GCACTGTGAC | AGAAGGTATT | ATTTCAGGTC | TGAACAGAAA | 12480 |
| 15 | TGTTCCTATT | GATTTCGATA | AAGATAATAA | ATATGATATG | TTGATGAAAG | CTTTCCAAAT | 12540 |
| | TGATGCATCA | GTAAATCCAG | GTAACTCGGG | TGGTGCTGTC | GTCAATAGAG | AAGGAAAATT | 12600 |
| io | AATAGGTGTA | GTTGCAGCTA | AAATTAGTAT | GCCAAACGTT | GAAAnTATGT | CATTTGCA | 12658 |
| | (0) TVD05:45 | | | _ | | | |

(2) INFORMATION FOR SEQ ID NO: 128:

⁽i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6048 base pairs

(C) STRANDEDNESS: double (D) TOPOLOGY: linear

| (xi) S | SEQUENCE DES | CRIPTION: S | SEQ ID NO: 1 | 128: | | |
|------------|---|--|---|---|---|---|
| TGAAATnGAA | TAGTACTATT | GCAAGTGTAA | AGAGGTTAAT | TTTTGCCnCA | CGCGGGACTT | 60 |
| AAAAAGGCAA | CCACTGGTTG | TGACATATCC | TTATTTACAT | TTATAAATAT | AAGGAGGAGG | 120 |
| TAGTAGTGAA | AGACTTATTG | CAAGCACAGC | AAAAGCTTAT | ACCGGATCTC | ATAGATAAAA | 180 |
| TGTATAAACG | TTTTTCTATT | CTTACTACTA | TCTCAAAAAA | TCAGCCTGTC | GGACGTCGAA | 240 |
| GTTTAAGCGA | ACATATGGAT | ATGACTGAAC | GTGTACTGCG | TTCTGAAACA | GATATGCTTA | 300 |
| AGAAACAAGA | TTTGATAAAA | GTTAAGCCTA | CCGGAATGGA | AATTACAGCT | GAAGGTGAGC | 360 |
| AACTGATTTC | GCAATTGAAA | GGTTACTTTG | ATATCTATGC | AGATGATAAT | CGTCTGTCAG | 420 |
| AAGGTATTAA | GAATAAATTT | CAAATTAAGG | AAGTTCATGT | TGTTCCTGGT | GATGCTGATA | 480 |
| ATAGTCAATC | TGTTAAAACA | GAATTAGGTA | GACAAGCAGG | TCAATTACTT | GAAGGCATAT | 540 |
| TACAAGAAGA | CGCGATAGTT | GCTGTAACTG | GCGGATCCAC | GATGGCATGT | GTTAGTGAAG | 600 |
| CAATTCATTT | ATTACCATAT | AATGTATTCT | TCGTACCAGC | CAGAGGTGGA | CTAGGCGAAA | 660 |
| ATGTTGTCTT | TCAGGCAAAC | ACAATTGCAG | CCAGTATGGc | aCAACAAGCT | GGCGGTTATT | 720 |
| ATACGACGAT | GTATGTACCT | GATAATGTCA | GTGAAaCAAC | ATATAATACA | TTGTTGTTAG | 780 |
| AGCCATCAGT | CATAAACACT | TTAGACAAAA | TTAAACAAGC | AAACGTTATA | TTACACGGCA | 840 |
| TTGGTGATGC | GCTGAAGATG | GCGCATCGAC | GTCAATCACC | TGAAAAGGTC | ATTGAACAAC | 900 |
| TTCAACATCA | TCAAGCTGTC | GGAGAGGCAT | TTGGTTATTA | TTTTGATACA | CAAGGTCAAA | 960 |
| TTGTCCATAA | GGTTAAAACA | ATTGGACTTC | AATTAGAAGA | CCTTGAATCA | AAAGACTTTA | 1020 |
| TTTTTGCAGT | TGCAGGAGGC | AAATCGAAAG | GTGAAGCAAT | TAAAGCATAC | TTGACGATTG | 1080 |
| CACCCAAGAA | TACAGTGTTA | ATCACTGATG | _AAGCCGCAGC | AAAGATAATA | _CTTGAATAAG: | 1140 |
| AGATAAAAAG | TTTAATACTT | TTTAAATATC | ATTTTAAAGG | AGGCCATTAT | AATGGCAGTA | 1200 |
| AAAGTAGCAA | TTAATGGTTT | TGGTAGAATT | GGTCGTTTAG | CATTCAGAAG | AATTCAAGAA | 1260 |
| GTAGAAGGTC | TTGAAGTTGT | AGCAGTAAAC | GACTTAACAG | ATGACGACAT | GTTAGCGCAT | 1320 |
| TTATTAAAAT | ATGACACTAT | GCAAGGTCGT | TTCACAGGTG | AAGTAGAGGT | AGTTGATGGT | 1380 |
| GGTTTCCGCG | TAAATGGTAA | AGAAGTTAAA | TCATTCAGTG | AACCAGATGO | AAGCAAATTA | 1440 |
| CCTTGGAAAG | ACTTAAATAT | CGATGTAGTA | TTAGAATGTA | CTGGTTTCTA | CACTGATAAA | 1500 |
| GATAAAGCAC | AAGCTCATAT | TGAAGCAGGC | GCTAAAAAAG | TATTAATCTC | AGCACCAGCT | 1560 |
| | TGAAATNGAA AAAAAGGCAA TAGTAGTGAA TGTATAAACG GTTTAAGCGA AGAAACAAGA AACTGATTTC AAGGTATTAA ATAGTCAATC TACAAGAAGA CAATTCATTT ATGTTGTCTT ATACGACGAT TTGGTGATGC TTCAACATCA TTGTCCATAA TTTTTGCAGT CACCCAAGAA AGATAAAAAG AAAGTAGCAA GTAGAAGGTC TTATTAAAAT GGTTTCCGCG CCTTGGAAAG | TGAAATNGAA TAGTACTATT AAAAAGGCAA CCACTGGTTG TAGTAGTGAA AGACTTATTG TGTATAAACG TTTTTCTATT GTTTAAGCGA ACATATGGAT AGAAACAAGA TTTGATAAAA AACTGATTTC GCAATTGAAA AAGGTATTAA GAATAAATTT ATAGTCAATC TGTTAAAACA TACAAGAAGA CGCGATAGTT CAATTCATTT ATTACCATAT ATGTTGTCTT TCAGGCAAAC ATACGACGAT GTATGTACCT AGCCATCAGT CATAAACACT TTGGTGATGC GCTGAAGATG TTCAACATCA TCAAGCTGTC TTGTCCATAA GGTTAAAACA TTTTTGCAGT TGCAGGAGGC CACCCAAGAA TACAGTGTTA AGATAAAAAG TTTAATACTT AAAGTAGCAA TTAATGGTTT GTAGGAAGGTC TTGAAGTTGT TTATTAAAAT ATGACACTAT GGTTTCCGCG TAAATGGTAA CCTTGGAAAG ACTTAAATAT | TGAAATIGAA TAGTACTATT GCAAGTGTAA AAAAAGGCAA CCACTGGTTG TGACATATCC TAGTAGTGAA AGACTTATTG CAAGCACAGC TGTATAAACG TTTTTCTATT CTTACTACTA GTTTAAGCGA ACATATGGAT ATGACTGAAC AGAAACAAGA TTTGATAAAA GTTAAGCCTA AACTGATTC GCAATTGAAA GGTTACTTTG AAGGTATTAA GAATAAATTT CAAATTAAGG ATAGTCAATC TGTTAAAACA GAATTAGGTA TACAAGAAGA CGCGATAGTT GCTGTAACTG CAATTCATTT ATTACCATAT AATGTATTCT ATGTTGTCTT TCAGGCAAAC ACAATTGCAG ATACGACGAT GTATAAACAT TTAGACAAAA TTGGTGATGC GCTGAAGATG GCGCATCGAC TTCAACATCA TCAAGCTGTC GGAGAGGCAT TTGTCCATAA GGTTAAAACA ATTGGACTTC TTTTTGCAGT TGCAGGAGGC AAATCGAAGG CACCCAAGAA TACAGTGTTA ATCACTGATG AGATAAAAAG TTTAATACTT TTTAAATATC AAAGTAGCAA TTAATGGTTT TGGTAGAATT GTAGAAGGTC TTGAAGTTGT AGCAGTAAAC TTATTAAAAT ATGACACTAT GCAAGGTCGT GGTTTCCGCG TAAATGGTAA AGAAGTTAAA | TGAAATHGAA TAGTACTATT GCAAGTGTAA AGAGGTTAAT AAAAAGGCAA CCACTGGTTG TGACATATCC TTATTTACAT TAGTAGTGAA AGACTTATTG CAAGCACAGC AAAAGCTTAT TGTATAAACG TTTTCTATT CTTACTACTA TCTCAAAAAA GTTTAAGCGA ACATATGGAT ATGACTGAAC GTGTACTGCG AGAAACAAGA TTTGATAAAA GTTAAGCCTA CCGGAATGGA AACTGATTTC GCAATTGAAA GGTTACTTTG ATATCTATGC AAGGTATTAA GAATAAATTT CAAATTAAGG AAGTTCATGT ATAGTCAATC TGTTAAAACA GAATTAAGG AAGTTCATGT ATGATGATTT ATTACCATAT AATGTATTCT TCGTACCAGC ATGATGTCTT TCAGGCAAAC ACAATTGCAG CCAGTATGGC ATGCTGATC GTATAAACAC GAATAGTCA GTGAAACAAC AGCCATCAGT GTATGTACCT GATAATGTCA GTGAAACAAC TTGGTGATGC GCTGAAGATG GCGCATCGAC GTCAATCACC TTCAACATCA TCAAGCTGTC GGAGAGGCAT TTGGTTATTA TTGTCCATAA GGTTAAAACAA ATTGGACTTC AATTAGAAGA TTTTTTGCAGT TGCAGGAGGC AAATCGAAG GTGAAGCAAT TTTTTTGCAGT TGCAGGAGGC AAATCGAAG GTGAAGCAAT CACCCAAGAA TACAGTGTTA ATCACTGATG AAGCCGCAGC AGATAAAAAG TTTAATACTT TTTAAATATC ATTTTAAAGG AAAGTAGCAA TTAATGGTTT TGGTAGAATC GACTTTAACAGG TTATTAAAAAA TTAATGGTTT TGGTAGAATC GACTTTAACAGG TTATTAAAAAA TTAATGGTTT TGGTAGAATC GACTTTAACAGG TTATTAAAAAA ATGAACTAT GGTAGAATACAGG TTATTAAAAAA ATGAACTAT GGTAGAATACAGG TTATTAAAAAAAA TTAATGGTTT TGGTAGAATT GGTCGTTTAGG GGTTTCCGCG TAAATGGTAA AGAAGTTAAA TCATTCAGTGG GGTTTCCGCG TAAATGGTAA AGAAGTTAAA TCATTCAGTGG CCCTTGGAAAG ACTTAAAATAT CGATGTAGAA TCATTCAGTGAGAACAAAA TTAAGAATGTAA AGAAGTTAAAA TCATTCAGTGAGAAAAAAAAAG TTTAAAATATA ATGAACTAAA TCATTCAGTGAAAAAAAAAA | AAAAAGGCAA CCACTGGTTG TGACATATCC TTATTTACAT TTATAAATAT TAGTAGTGAA AGACTTATTG CAAGCACAGC AAAAGCTTAT ACCGGATCTC TGTATAAACG TTTTCTATT CTTACTACTA TCTCAAAAAA TCAGCCTGTC GTTTAAGCGA ACATATGGAT ATGACTGAAC GTGTACTGCG TTCTGAAACA AGAAACAAGA TTTGATAAAA GTTAAGCCTA CCGGAATGGA AATTACAGCT AACTGATTTC GCAATTGAAA GGTTACTTG ATATCATGC AGATGATAAT AAAGGTATTAA GAATAAATTT CAAATTAAGG AAGTTCATGT TGTTCCTGGT ATAGTCAATC TGTTAAAACA GAATTAAGG AAGTTCATGT TGTTCCTGGT TACAAGAAGA CGCGATAGTT GCTGTAACTG GCGGATCCAC GATGGCATGT CAATTCATTT ATTACCATAT AATGTATTCT TCGTACCAGC CAGAGGTGGA ATGTTGTCTT TCAGGCAAAC ACAATTGCAG CCAGTATGGC ACAACAAGCT ATACGACGAT GTATGACCT GATAATGTCA GTGAAACAAC ATATAATACA AGCCATCAGT CATAAACACT TTAGACAAAA TTAAACAAGC AAACGTTATA TTGGTGATGC GCTGAAGATG GCGCATCGAC GTCAATCACC TGAAAAGGTC TTCAACATCA TCAAGCTGTC GGAGAGGCAT TTGGTTATTA TTTTGATACA TTTTTGCAGT TGCAGGAGGC AAATCGAAAG GTGAAGCAAT TAAAGCATCA TTTTTTGCAGT TGCAGGAGGC AAATCGAAAG GTGAAGCAAT TAAAGCATCA CACCCAAGAA TACAGTGTTA ATCACTGATG AAGCCGCAGC AAAGATAATA AGATAAAAAG TTTAATACTT TTTAAATATC ATTTTAAAGA AGGCCATTAT AAAGTAGCAA TTAATGCTT TGGTAGAAT GGTCGTTTAG CATTCAGAAG GTGAAAGAAC TTAATGCTT TGGTAGAAT GGTCGTTTAG CATTCAGAAG GTAGAAGGAC TTGAAGTTAT AGCCGCAGC AAAGATAATA AAAGTAACAAG TTTAATACTT TTTAAATATC ATTTTAAAGA AGCCCATTAT AAAGCATAAA AAAGTAAAAAG TTTAATACTT TTTAAATATC ATTTTAAAGA AGCCCATTAT AAAGCATAAA AAAGTAAAAAA TTAAATGTTT TGGTAGAAT GGTCGTTTAG CATTCAGAAG GTAGAAGGTC TTGAAGTTGT AGCAGTAAAC GACTTAACAG ATGACGACAT TTATTAAAAA ATGACACTAT GCAAGGTCGT TTCACAGGTG AAGCACAT TTATTAAAAA ATGACACTAT GCAAGGTCGT TTCACAGGTG AAGTAGAGGT TTATTAAAAAA ATGACACTAT GCAAGGTCGT TTCACAGGTG AACCAGATGC CCTTGGAAAG ACTTAAATAT CGATGTAAA TCATTCAGTG AACCAGATGC CCTTGGAAAG ACTTAAATAT CGATGTAAA TCATTCAGTG AACCAGATGC CCTTGGAAAG ACTTAAATAT CGATGTAAAA TCATTCAGTG AACCAGATGC CCTTGGAAAG ACTTAAATAT CGATGTATA CAGTTTAAAAAG ACTTAAATAT CGATGTATAA CGATTAAAA CCAGATGC CCTTGGAAAG ACTTAAATAT CGATGTATA CAGTTGTA TTAAAGAGA CCTTGATTCAACAG TTCACAGGTG AACCAGATGC | TGAAATHGAA TAGTACTATT GCAAGTGTAA AGAGGTTAAT TITTGCCHCA CGCGGGACTT AAAAAGGCAA CCACTGGTTG TGACATATCC TTATTTACAT TTATAAATAT AAGGAGGGG TAGTAGTGAA AGACTTATTG CAAGCACAGC AAAAGCTTAT ACCGGATCTC ATAGATAAAA TGTATAAACG TTTTTCTATT CTTACTACTA TCTCAAAAAA TCAGCCTGTC GGACGTCGAA GTTTAAGCGA ACATATGGAT ATGACTGAAC GTGTACTGCG TTCTGAACACA GATATGCTTA AGAAACAAGA TTTGATAAAA GTTAAGCCTA CCGGAATGGA AATTACAGCT GAAGGTGAGC AACTGATTCA GCAATTGAAA GGTTACTTG ATATCTATGC AGATGATAATA CGTCTGTCAG AAGGTATTAA GAATAAAATT CAAATTAAGG AAGTTCATGT TGTTCCTGGT GATGCTGATA ATAGTCAATC TGTTAAAACA GAATTAGGTA GACAAGCAGG TCAATTACTT GAAGGCATAT TACAAGAAGA CGCGATAGTT GCTGTAACTG GCGGATCCAC GATGGCATGT GTTAGTGAAG AATGTCAATC TGTTAAAACA GAATTACGT GCGGATCCAC GATGGCATGT GTTAGTGAAG CAATTCATTT ATTACCATAT AATGTTATCT TCGTACCAGC CAGAGGTGGA CTAGGCGAAA ATGTTGTCTT TCAGGCAAAC ACAATTGCAG CCAGTATGGC ACAACAAGCT GCCGGTTATT ATACGACGAT GTATGTACCT GATAATGTCA GTGAAACAAC ATATAATACA TTGTTGTTAG AGCCATCAGT CATAAACACT TTAGACAAAA TTAAACAAC AAACGTTAAT TTACACGGCA TTGGTGATGC GCTGAAACAC TTAGACAAAA TTAAACAAGC AAACGTTATA TTACACGGCA TTGGTGATGC GCTGAAACAC ATGGACACA GTCAATCACC TGAAAAGGTC ATTGAACAAC TTCAACATCA TCAAGCTGTC GGAGAGGCAT TTGGTTATTA TTTTGATCAC CAAGGTCAAA TTTTTTGCATT TGCAGGAGGC AAATCGAAG GTCAATCACC TGAAAAGGTC ATTGAACAAC TTTTTTTGCATT TGCAGGAGGC AAATCGAAAG CTTGAATCAA AAAGACTTTA TTTTTTGCATT TGCAGGAGGC AAATCGAAAG GTGAAGCAAT TAAAGCATAC TTGACCATAA AAAGTACAAA TACAGTGTTA ATCACTGAAG AGCCGCAGC AAAGATAATA CTTGAATAAG AGATAAAAAG TTTAATACTT TTTAAATATC ATTTTAAAGG AGGCCATTAT AATGGCAGTA AAAATGAGAA TACAGTGTTA ACCACTGAAA GACTTAAC GATCAACAG TTGAACAGAC TTAAAAAGAC TTGAACAGAC AAGCACAATTA CCATGATAAA TTAACACCAGATGC TTAAACAGC TTGAACAGAC AAGCCAACAC TTAAAAAAC CAAGACTAAA TTAACACTATA |

| | ACAGTTGTTT | CAGGTGCTTC | ATGTACTACA | AACTCATTAG | CACCAGTTGC | TAAAGTTTTA | 1680 |
|----|------------|------------|------------|------------|------------|------------|------|
| | AACGATGACT | TTGGTTTAGT | TGAAGGTTTA | ATGACTACAA | TTCACGCTTA | CACAGGTGAT | 1740 |
| 5 | CAAAATACAC | AAGACGCACC | TCACAGAAAA | GGTGACAAAC | GTCGTGCTCG | TGCAGCGGCA | 1800 |
| | GAAAACATCA | TCCCTAACTC | AACAGGTGCT | GCTAAAGCTA | TCGGTAAAGT | TATTCCTGAA | 1860 |
| 10 | ATCGATGGTA | AATTAGATGG | TGGTGCACAA | CGTGTTCCTG | TAGCTACAGG | TTCATTAACT | 1920 |
| | GAATTAACAG | TAGTATTAGA | AAAACAAGAC | GTAACAGTTG | AACAAGTTAA | CGAAGCTATG | 1980 |
| | AAAAATGCTT | CAAACGAATC | ATTCGGTtAC | ACTGAAGACG | AAATCGTTTC | TTCAGACGTT | 2040 |
| 15 | GTAGGTATGA | CTTACGGTTC | ATTATTCGAC | GCTACACAAA | CTCGTGTAAT | GTCAGTTGGC | 2100 |
| | GACCGTCAAT | TAGTTAAAGT | TGCAGCTTGG | TATGATAACG | AAATGTCATA | TACTGCACAA | 2160 |
| | TTAGTTCGTA | CATTAGCATA | CTTAGCTGAA | CTTTCTAAAT | AATTTTAGTA | TAGTTTTTAT | 2220 |
| 20 | TCAAATACGC | TAGTGCTCAG | AACTATTTAG | CATTAATTAA | AGCTTATGAG | TAAGCGGGGA | 2280 |
| | GCACAAACGC | TTCTCCGCTT | ATTTTTATAT | AAAATTTCCT | AATTACAAGG | AGGAAACACC | 2340 |
| | ATGGCTAAAA | AAATTGTTTC | TGATTTAGAT | CTTAAAGGTA | AAACAGTCCT | AGTACGTGCT | 2400 |
| 25 | GATTTTAACG | TACCTTTAAA | AGACGGTGAA | ATTACTAATG | ACAACCGTAT | CGTTCAAGCT | 2460 |
| | TTACCTACAA | TTCAATACAT | CATCGAACAA | GGTGGTAAAA | TCGTACTATT | TTCACATTTA | 2520 |
| 30 | GGTAAAGTGA | AAGAAGAAAG | TGATAAAGCA | AAATTAACTT | TACGTCCAGT | TGCTGAAGAC | 2580 |
| | TTATCTAAGA | AATTAGATAA | AGAAGTTGTT | TTCGTACCAG | AAACACGCGG | CGAAAAACTT | 2640 |
| | GAAGCTGCTA | TTAAAGACCT | TAAAGAAGGC | GACGTATTAT | TAGTTGAAAA | TACACGTTAT | 2700 |
| 35 | GAAGATTTAG | ACGGTAAAAA | AGAATCTAAA | AATGATCCAG | AATTAGGTAA | ATACTGGGCA | 2760 |
| | TCTTTAGGTG | ATGTGTTTGT | AAATGATGCT | TTTGGTACTG | CGCATCGTGA | GCATGCATCT | 2820 |
| | AATĢTTGGTA | TTTCTACACA | TTTAGAAACT | GCAGCTGGAT | TCTTAATGGA | TAAAGAAATT | 2880 |
| 40 | AAGTTTATTG | GCGGCGTAGT | TAACGATCCA | CATAAACCAG | TTGTTGCTAT | TTTAGGTGGA | 2940 |
| | GCAAAAGTAT | CTGACAAAAT | TAATGTCATC | AAAAACTTAG | TTAACATAGC | TGATAAAATT | 3000 |
| | ATCATCGGCG | GAGGTATGGC | TTATACTTTC | TTAAAAGCGC | AAGGTAAAGA | AATTGGTATT | 3060 |
| 45 | TCATTATTAG | AAGAAGATAA | AATCGACTTC | GCAAAAGATT | TATTAGAAAA | ACATGGTGAT | 3120 |
| | AAAATTGTAT | TACCAGTAGA | CACTAAAGTT | GCTAAAGAAT | TTTCTAATGA | TGCCAAAATC | 3180 |
| 50 | ACTGTAGTAC | CATCTGATTC | AATTCCAGCA | GACCAAGAAG | GTATGGATAT | TGGACCAAAC | 324 |
| | ACTGTAAAAT | TATTTGCAGA | TGAATTAGAA | GGTGCGCACA | CTGTTGTATG | GAATGGACCT | 330 |
| | ATGGGTGTAT | TCGAGTTCAG | TAACTTTGCA | CAAGGTACAA | TTGGTGTATG | TAAAGCAATT | 336 |

| | TCTTTAGGTT | TTGAAAATGA | CTTCACTCAT | ATTTCAACTG | GTGGCGGCGC | GTCATTAGAG | 3480 |
|----|------------|------------|------------|------------|------------|------------|------|
| | TACCTAGAAG | GTAAAGAATT | GCCTGGTATC | AAAGCAATCA | ATAATAAATA | ATAAAGTGAT | 3540 |
| 5 | AGTTTAAAGT | GATGTGGCAT | GTTTGTTTAA | CATTGTTACG | GGAAAACAGT | CACAAGATGA | 3600 |
| | CATCGTGTTT | CATCACTTTT | CAAAAATATT | TACAAAACAA | GGAGTGTCTT | TAATGAGAAC | 3660 |
| 10 | ACCAATTATA | GCTGGTAACT | GGAAAATGAA | CAAAACAGTA | CAAGAAGCAA | AAGatTCGTC | 3720 |
| 10 | AATACATTAC | CAACACTACC | AGATTCAAAA | GAAGTAGAAT | CAGTAATTTG | TGCACCAGCA | 3780 |
| | ATTCAATTAG | ATGCATTAAC | TACTGCAGTT | AAAGAAGGAA | AAGCACAAGG | TTTAGAAATC | 3840 |
| 15 | GGTGCTCAAA | ATACGTATTT | CGAAGATAAT | GGTGCGTTCA | CAGGTGAAAC | GTCTCCAGTT | 3900 |
| | GCATTAGCAG | ATTTAGGCGT | TAAATACGTT | GTTATCGGTC | ATTCTGAACG | TCGTGAATTA | 3960 |
| | TTCCACGAAA | CAGATGAAGA | AATTAACAAA | AAAGCGCACG | CTATTTTCAA | ACATGGAATG | 4020 |
| 20 | ACTCCAATTA | TATGTGTTGG | TGAAACAGAC | GAAGAGCGTG | AAAGTGGTAA | AGCTAACGAT | 4080 |
| | GTTGTAGGTG | AGCAAGTTAA | GAAAGCTGTT | GCAGGTTTAT | CTGAAGATCA | ACTTAAATCA | 4140 |
| | GTTGTAATTG | CTTATGAACC | AATCTGGGCA | ATCGGAACTG | GTAAATCATC | AACATCTGAA | 4200 |
| 25 | GATGCAAATG | AAATGTGTGC | ATTTGTACGT | CAAACTATTG | CTGACTTATC | AAGCAAAGAA | 4260 |
| | GTATCAGAAG | CAACTCGTAT | TCAATATGGT | GGTAGTGTTA | AACCTAACAA | CATTAAAGAA | 4320 |
| 30 | TACATGGCAC | AAACTGATAT | TGATGGGGCA | TTAGTAGGTG | GCGCATCACT | TAAAGTTGAA | 4380 |
| | GATTTCGTAC | AATTGTTAGA | AGGTGCAAAA | TAATCATGGC | TAAGAAACCa | ACTGCGTTAA | 4440 |
| | TTATTTTAGA | TGGTTTTGCG | AACCGCGAAA | GCGAACATGG | TAATGCGGTA | AAATTAGCAA | 4500 |
| 35 | ACAAGCCTAA | TTTTGATCGT | TATTACAACA | AATATCCAAC | GACTCAAATC | GAAGCGAGTG | 4560 |
| | GCTTAGATGT | TGGACTACCT | GAAGGACAAA | TGGGTAACTC | AGAAGTTGGT | CATATGAATA | 4620 |
| | TCGGTGCAGG | ACGTATCGTT | TATCAAAGTT | TAACTCGAAT | CAATAAATCA | ATTGAAGACG | 4680 |
| 40 | GTGATTTCTT | TGAAAATGAT | GTTTTAAATA | ATGCAATTGC | ACACGTGAAT | TCACATGATT | 4740 |
| | CAGCGTTACA | CATCTTTGGT | TTATTGTCTG | ACGGTGGTGT | ACACAGTCAT | TACAAACATT | 4800 |
| | TATTTGCTTT | GTTAGAACTT | GCTAAAAAAC | AAGGTGTTGA | AAAAGTTTAC | GTACACGCAT | 4860 |
| 45 | TTTTAGATGG | CCGTGACGTA | GATCAAAAAT | CCGCTTTGAA | ATACATCGAA | GAGACTGAAG | 4920 |
| | CTAAATTCAA | TGAATTAGGC | ATTGGTCAAT | TTGCATCTGT | GTCTGGTCGT | TATTATGCAA | 4980 |
| 50 | TGGATCGTGA | CAAACGTTGG | GAACGTGAAG | AAAAAGCTTA | CAATGCTATT | CGTAATTTTG | 5040 |
| | ATGCCCCAAC | TTATGCAACT | GCCAAAGAAG | GTGTAGAAGC | AAGCTATAAT | GAGGGCTTAA | 5100 |
| | CTGACGAATT | CGTAGTACCA | TTCATCGTTG | AGAATCAAAA | TGACGGTGTT | AATGATGGAG | 5160 |

| | CGAACAGAGC ATTCGAAGGC TTTAAAGTTG AACAAGTTAA AGACTTATTC TATGCAACAT | 5280 |
|----|---|------------|
| _ | TCACTAAGTA TAATGACAAT ATCGATGCGG CTATCGTCTT CGAAAAAGTT GATTTAAATA | 5340 |
| 5 | ATACAATTGG TGAAATTGCA CAAAATAACA ATTTAACTCA ATTACGTATT GCAGAAACTG | 5400 |
| | AAAAATACCC TCACGTTACT TACTTTATGA GTGGTGGACG TAACGAGGAA TTTAAAGGTG | 5460 |
| 10 | AACGCCGTCG TTTAATTGAT TCACCTAAAG TTGCAACGTA TGACTTGAAA CCAGAAATGA | 5520 |
| | GTGCTTATGA AGTTAAAGAT GCATTATTAG AAGAGTTAAA TAAAGGTGAC TTGGACTTAA | 5580 |
| | TTATTTTAAA CTTTGCTAAC CCTGATATGG TTGGACATAG TGGTATGCTT GAGCCGACAA | 5640 |
| 15 | TCAAAGCAAT CGAAGCGGTT GATGAATGTT TAGGAGAAGT GGTTGATAAG ATTTTAGACA | 5700 |
| | TGGACGGTTA TGCAATTATT ACTGCTGACC ATGGTAACTC TGATCAAGTA TTGACGGATG | 5760 |
| | ATGATCAACC AATGACTACG CAWACAACGA ACCCAGTACC AGTGATTGTA ACAAAAGAAG | 5820 |
| 20 | GCGTTACACT TAGAGAAACT GGTCGCTTAG GTGACTTAGC ACCTACATTA TTAGATTTAT | 5880 |
| | TAAATGTAGA ACAACCTGAA GATATGACAG GTGAATCTTT AATTAAACAC TAATATTGTA | 5940 |
| | AAAGATGTTA AGTAAACGCT TAATGACACT TATTTTTTGA AAATAATAGT AATATCNTTT | 6000 |
| ?5 | TGTTAAATGA AAGAATAAAG CTATAATAAT TATAGAATAA CTATTTAN | 6048 |
| | (2) INFORMATION FOR SEQ ID NO: 129: | • |
| | (i) SEQUENCE CHARACTERISTICS: | |
| 30 | (A) LENGTH: 5602 base pairs (B) TYPE: nucleic acid | • |
| | (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 25 | | |
| 35 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 129: | ~ " |
| | AAAGAAGTGC AAGATATCAT CGCATTAATT AAGTCGTTAC AAAGTGTAAT TGTAGACATC | 60 |
| 40 | GCTTCCAATA ATGTTGATAC AATTATGCCT GGTTATACTC ATTTACAGCG TGCACAGCCA | 120 |
| | ATTTCATTTG CACATCATAT TATGACTTAT TTTTGGATGT TACAACGAGA CCAACAACGA | 180 |
| | TTTGAAGATA GTTTAAAACG AATCGATATT AATCCTTTAG GTGCAGCAGC CTTAAGTGGT | 240 |
| 45 | ACCACATACC CTATCGATAG ACACGAGACA ACAGCATTGT TGAACTTTGG CAGTCTCTAT | 300 |
| | GAGAATAGCC TAGATGCTGT TAGTGACAGA GACTATATTA TTGAAACATT GCATAATATT | 360 |
| | TCTTTAACGA TGGTTCACTT ATCACGCTTT GCAGAGGAAA TTATTTTCTG GTCCACAGAC | 420 |
| 50 | GAAGCTAAAT TCATTACATT ATCAGATGCA TTTTCAACTG GCTCATCTAT TATGCCACAA | 480 |
| | | |

55

540

AAGAAAAATC CTGATATGGC AGAATTAATT AGAGGTAAAG TTGGTCGAAC GACTGGTCAT

| | GAAGATAAAG | AAGGTTTATT | CGATGCTGTC | CATACAATTA | AAGGTTCTTT | ACGTATTTTC | 660 | |
|----------------|------------|------------|------------|-------------------------|--------------------|--------------|------|---|
| _ | GAAGGTATGA | TTCAAACGAT | GACAATTAAT | AAAGAACGAC | TCAATCAAAC | TGTTAAAGAA | 720 | |
| 5 | GATTTTTCAA | ATGCAACGGA | ACTAGCAGAT | TATTTAGTAA | CTAAAAATAT | TCCATTTAGA | 780 | |
| | ACTGCACATG | AAATTGTAGG | AAAAATCGTC | TTAGAATGTA | TACAACAAGG | TCATTATTTA | 840 | |
| 10 | TTAGATGTTC | CTTTAGCAAC | ATATCAACAA | CATCATTCTA | GTATTGATGC | CGATATTTAC | 900 | |
| | GATTATTTGC | AGCCTGAAAA | TTGTTTAAAA | CGACGTCAAA | GTTACGGTTC | AACAGGTCAA | 960 | |
| | TCATCGGTCA | AACAACAACT | TGATGTTGCT | AAACAATTAC | TATCACAATA | AATACGTTAA | 1020 | |
| 15 | TCTACCTACC | CACAATGTCT | ATTAAAATTA | CATTGTGGGT | ATTTTAATGC | TCTCTTCGTC | 1080 | |
| • | TTGTTGAACA | TCACATTITT | AAGATTCCTA | AAATGTTTGA | TAATTCTTTT | AAATTTATAT | 1140 | |
| | TACAAAAATG | TTATAAATTG | TAAAAGAAAT | GTGTAAAGCG | TTTTCACAAG | CAGGTTTTTG | 1200 | |
| 20 | TAGTATTTTA | AAATTGTTAG | ACTACAAATA | AAGAGATGAA | AGGATAAAGA | CTATGACTAA | 1260 | |
| | CTCTTCGAAA | AGCTTCACTA | AATTTATGGC | TGCTTCTGCT | GTTTTTACTA | TGGGATTTTT | 1320 | |
| | ATCAGTACCT | ACTGCTGGCG | CTGAACAAAC | AAATCAAATT | GCAAATAAAC | CTCAGGCTAT | 1380 | |
| 25 | TCAATGGCAT | ACAAATTTAA | CGAATGAGCG | ATTCACTACT | ATCGCACATC | GTGGCGCAAG | 1440 | |
| | TGGCTATGCA | CCCGAGCATA | CGTTTCAAGC | ATATGATAAG | AGTCATAATG | AGTTAAAAGC | 1500 | |
| 30 | ATCTTATATC | GAAATTGATT | TACAACGTAC | CAAAGATGGC | CATTTAGTTG | CTATGCATGA | 1560 | |
| | TGAAACTGTT | AACCGTACAA | CAAATGGACA | CGGTAAAGTT | GAGGATTATA | CCCTTGATGA | 1620 | |
| | ATTAAAACAG | TTAGATGCAG | GAAGTTGGTT | ТААТААААА | TATCCAAAAT | ACGCAAGAGC | 1680 | |
| 35 | AAGTTATAAA | AATGCTAAAG | TACCCACTTT | AGATGAAATT | TTAGAACGTT | ATGGCCCGAA | 1740 | |
| | TGCAAACTAT | TATATTGAAA | CAAAGTCACC | TGATGTATAC | CCAGGAATGG | AAGAACAATT . | 1800 | |
| | ATTAGCTTCA | TTGAAAAAGC | ATCACCTTTT | AAATAACAAT | AAATTAAAA | ATGGACATGT | 1860 | |
| 40 | AATGATTCAA | TCATTTTCTG | ACGAAAGTTT | AAAGAAAATT | CATCGTCAAA | ATAAGCATGT | 1920 | |
| | GCCATTAGTA | AAATTAGTTG | ATAAAGGTGA | ACTACAACAA | TTTAACGACC | AACGCTTAAA | 1980 | |
| | AGAGATACGC | TCTTATGCGA | TTGGATTAGG | TCCTGATTAT | ACAGATTTAA | CTGAACAAAA | 2040 | • |
| 45 | TACCCATCAT | TTAAAAGACT | TAGGATTTAT | AGTACATCCT | TATACAGTGA | ATGAAAAAGC | 2100 | ` |
| | TGATATGTTA | CGATTAAATA | AATATGGCGT | TGATGGTGTC | TTTACAAATT | TCGCTGATAA | 2160 | |
| 50 | ATATAAAGAA | GTCATTAAGT | agtaatgtta | AACTAGAAAA | CATAAATACA | AAAATATAGC | 2220 | |
| - - | TATTACTATA | AAAAACAGCA | GTAAGATATT | TCCAAATTGA | AATTATCCTA | CTGCTGTCTT | 2280 | |
| | TTTCCCACTC | CCACACAAAT | CATATTTCC | <i>ሮ እ እ እ እ</i> ጥጥጥ እጥ | ポポインボボインボイン | CACCCCAACT | 2240 | |

| | TTGTCTGTAG | AAATTGAGGA | GCTAATTTCT | CTGTGTCGGG | GCTCCACCCC | AACTTGCACA | 2460 |
|----|------------|------------|------------|------------|------------|-------------|------|
| | CTATTGTAAG | CTGACTTTCC | GCCAGCCTCT | GTGTTGGGGC | CCCGCCAACT | TGCACACTAT | 2520 |
| 5 | TGTAAGCTGA | CTTTCCACCA | GCCTCTGTGT | TGGGGCCCCG | ACTATTTTTG | AAAAGAGCGT | 2580 |
| | GTTACACGGG | CATTGTTTTA | CAGTCAACTA | CTGCTAAAAT | AAAATTAACG | AGCTTAGGGC | 2640 |
| 10 | TTTGTTTTCT | GTCCCAAGCT | CGTTAAATCA | CATATGAȚAA | TTAATTATGC | CCAACCACGA | 2700 |
| | TATCTAGCTG | CTTCTGCTGT | ACGTTTAATA | CCTATGATAT | ATGCTGCAAG | TCTCATATCT | 2760 |
| | ATTTTTCGGT | TTTGAGACAA | TTCGTAAATC | GTATCAAATG | CCGCTTCTAA | TTTTTCACGT | 2820 |
| 15 | AGCTTTTCAT | TAACTTCTTC | TTCAGACCAA | TAATAACCTT | GATTATTTTG | TACCCATTCG | 2880 |
| | AAGTAAGAAA | CCGTLACACC | ACCAGCACTT | GCTAATACGT | CTGGAACTAA | TAATATACCA | 2940 |
| | CGTTCAGTTA | AAATACGTGT | TGCTTCTGGT | GTTGTAGGTC | CATTAGCAGC | TTCAACAACG | 3000 |
| 20 | ATACTAGCTT | TAATATCATG | TGCATTGTCT | TCTGTAATTT | GGTTTGAAAT | AGCCGCTGGT | 3060 |
| | ACTAAAATGT | CACAATCTAA | TTCAAACAAT | TCTTTATTTG | AGATTGTTTC | TTCAAATAAA | 3120 |
| | TTTGTTACCG | TACCAAAACT | ATCACGACGG | TCTAATAAAT | AATCTATATC | TAAGCCATTT | 3180 |
| 25 | GGATCGTGTA | ATGCACCGTA | AGCATCAGAG | ATACCTACAA | TTTTTGCACC | TAAATCATAT | 3240 |
| | AAGAATTTAG | CTAAGAAACT | TCCGGCATTA | CCGAAACCTT | GAATAACAAC | CTTGGCACCT | 3300 |
| 30 | TCAATTTGCA | TATTACGACG | TTTTGCAGCT | TGTTCAATTG | CAATAACTAC | ACCTAGTGCA | 3360 |
| | GTTGATCTGT | CGCGTCCATG | AGAACCACCC | AATACAATTG | GTTTACCTGT | GATGAAACCT. | 3420 |
| • | GGTGAATTAA | ATTTATCTAA | TGCACTATAT | TCATCCATCA | TCCAAGCCAT | AATTTGTGAG | 3480 |
| 35 | TTTGTAAATA | CATCTGGTGC | TGGAATATCT | TTGTTCGGAC | CTACGAATTG | TGAAATTGCT | 3540 |
| | CTTACATATC | CGCGTGATAA | ACGTTCAACT | TCATGAATGC | TCATTTGACG | TGGATCACAA | 3600 |
| | ACGATACCAC | CCTTACCACC | ACCGTATGGT | AAGTTTACAA | TGCCACATTT | CAAAGTCATC | 3660 |
| 40 | CACATTGATA | ATGCTTTTAC | TTCTTCTTCA | TCAACATCTG | GGTGGAAACG | CACGCCCCCT | 3720 |
| | TTTGTTGGTC | CAACAGCATC | ATTATGTTGC | GCACGGTAAC | CTGTGAATGT | TTTTACTGTG | 3780 |
| _ | CCATCATCCA | TTCGTACAGG | GATACGCACT | TGTAACATTC | TTAAAGGTTC | TTTAATTAAA | 3840 |
| 45 | TCGTACATTC | CTECGTCAAA | TCCCAATTTA | TGCAATGCTT | CTTTAATAAT | TCCTTGAGTA | 3900 |
| | GAAGTTACTA | AATTATTGTT | CTCAGTCATG | ATCCTTTTCG | CCTCTTCTTT | ACCTAATGAT | 3960 |
| 50 | TTCGCTTTCA | AACATATTGT | AACATAACGT | ATTCCTTTTT | AAAGCCCTTA | CAAACTGATT | 4020 |
| | GTTACAACTT | TTTGACATTA | TTGAAATACA | TGTCTTATTT | TTTCAAGTGC | AAGGTCCAAT | 4080 |
| | TCTTCTTTAG | TAATAATTAA | TGGTGGTGCA | AAACGAATGA | CAGTATCATG | CCLIMATOR | 4:40 |

| | ACACCTATAA | ACAAACCACG | TCCACGGACT | TCTTTAATTG | ATGGATGATC | AATTTGCTTT | 4260 |
|----|------------|------------|------------|------------|------------|------------|------|
| , | AATTGTTCTT | TAAAATAATC | TCCTAATTCT | AAAGAGCGGC | CTGGTAAATC | CTCATCAACG | 4320 |
| 5 | ATAACATCTA | ATGCAGCAAT | TGATGCAGCA | CAAGCAAGTG | GATTACCACC | AAATGTTGAA | 4380 |
| | CCATGTGAGC | CAGGTGTAAA | GACATCTAAT | ACTTCTTTAT | CTGCTAATAC | AACAGAAATT | 4440 |
| | GGGAAGACTC | CACCACCTAG | TGCTTTACCT | AAAATATAGA | CATCAGGTTT | TACATTATCC | 4500 |
| 0 | CAATCCGTAG | CAAATAATTT | ACCCGAACGA | CCTAATCCTG | CTTGGATTTC | GTCAGCAATA | 4560 |
| | AATAAGACAT | TATGTTCATC | ACATAATTCT | CTAATTGCTT | TCAAATATCC | TTCTGGCGGT | 4620 |
| 15 | ATATTTATAC | CCGCTTCACC | TTGAATTGGT | TCTACTAAAA | CTGCTGCAGT | ATTTTCATTA | 4680 |
| 3 | ATTGCAGCTT | TCAATGCATC | TACATCTCCA | AAATCAACTT | TTCTAAATCC | ATCTAATAAC | 4740 |
| | GGACCATAAC | CACGTTGGTA | TTCTGCTTCT | GAAGATAATG | AAACTGGCGC | CATTGTTCGA | 4800 |
| 20 | CCATGGAAGT | TACCATTAAA | TGCAATGATT | TCTGCTTTAT | TTGGCTCAAT | TCCTTTAACA | 4860 |
| | TCGTATGCCC | AGCGTCGTGC | TGCTTTCAAA | GCTGTTTCTA | CTGCTTCAGC | ACCTGTATTC | 4920 |
| | ATTGGTAAAG | CTTTATCTTT | ACCTGCCAGT | TTACAAATTT | TTTCGTACCA | TTCACCTAAG | 4980 |
| 25 | TTATCACTAT | GAAAAGCACG | TGAAACTAAA | GTCACTTTAT | CAGCTTGATC | TTTTAATGCT | 5040 |
| | TGAATAATTT | TCGGATGTCT | ATGACCTTGG | TTAACAGCGG | AATATGCAGA | TAACATATCC | 5100 |
| , | ATATATTTAT | TGCCTTCAGG | ATCTTTAACC | CATACCCCTT | CAGCTTcTGa | AATGaCAATT | 5160 |
| 30 | GGCAATGGTA | AATAATTATG | TGCTCCGTAA | TGATTTGTTA | ACTCAATAAT | TTTTTCAGAT | 5220 |
| | TTAGTCATCA | TATCTCCCCT | TTTCATCATT | TATAACTATT | ATACATGAAA | CATTATCCAA | 5280 |
| 25 | ATAATTACAT | TAGTTTTCAA | AGCAGATACT | TTTCCACCAA | AAAAGATGAA | ATAATCACTA | 5340 |
| 35 | AGTTTCATTA | AATTTGTCTA | TTTTGAAAAC | CCTTACATTT | ATAATGACAT | AATTACTTAA | 5400 |
| | ATGATTACAA | GCAAAAGAAT | TGATAATTTT | ACACTTAATC | AAAAGTATAT | TTTACTAAGA | 5460 |
| 40 | ATATTTTTAT | TTATAAATAT | TGAAAACCAC | TAACAAATTG | CATACACAAT | ATCATTAGTG | 5520 |
| | GTAACAGTTA | AACACTTATT | TATCTTTACG | GGGTAATGGG | TTAAAACCCT | ThCATTAAAA | 5580 |
| | TTGGATGnCC | ATAAAATTAG | GG | | | | 5602 |

(2) INFORMATION FOR SEQ ID NO: 130:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5924 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

 - (D) TOPOLOGY: linear

| | TAACCCCATT | TTACCTGGAA | AAATCgTTTG | CGATGCaATm | GCaTTtGaAT | ATAAATACAT | 60 |
|-----|--------------------|------------|------------|------------|------------|------------|------|
| | TTTACGTATa | GAATTATAAA | AgGTTTCATT | Caaatcttag | GGTCAAAAAT | GTTATAATAT | 120 |
| 5 | TTTTATGTCA | AATTTAAAAC | AGTAACACTT | ATTTACAAGG | TTGCAATATT | TTGAAGTAAT | 180 |
| | AAAGGAAGTG | TCGCGTATTT | TAACTTTTTC | AGAGCAAAAT | GCACTCGCGA | AAATAGATGA | 240 |
| 10 | TTTAATGAAT | ACTTATTGCA | ATCAATGTCC | AATCAAAACT | CGTCTGCGTA | AATTAGAGGG | 300 |
| ,,, | GAAAACGAAG | GCGCATCATT | TTTGTATCAA | TGAGTGTTCA | ATAGGGAAAG | AAATAAAACA | 360 |
| | ATTAGGAAAT | GAACTTCAAT | AGGAGGAAGT | CAAATGAAAA | TTATATCTAT | ATCAGAAACA | 420 |
| 15 | CCGAACCACA | ACACAATGAA | GATTACACTT | AGTGAAAGCA | GAGAAGGTAT | GACATCAGAT | 480 |
| | ACGTATACTA | AAGTTGATGA | TTCACAGCCA | GCATTTATTA | ATGACATCTT | AAAGGTTGAA | 540 |
| | GGCGTTAAAT | CAATTTTCCA | TGTTATGGAC | TTTATTTCAG | TAGATAAAGA | AAATGACGCA | 600 |
| 20 | AATTGGGAAA | CAGTATTGCC | AAAAGTAGAG | GCTGTATTCG | AATAAATTTT | TCATCAACTA | 660 |
| | GTATTCGGGG | GGAATAAAGT | ATATGGAAAT | TTTACGTATA | GAGCCAACAC | CAAGTCCAAA | 720 |
| | TACAATGAAA | GTTGTTTTGT | CATATACAAG | AGAAGACAAG | TTATCTAATA | CTTATAAAAA | 780 |
| 25 | AGTAGAAGAA | ACACAACCAA | GATTTATAAA | TCAGTTGTTA | TCTATAGATG | GTATCACTTC | 84.0 |
| | CATTTTTCAT | GTCATGAACT | TCTTAGCTGT | TGATAAGGCA | CCAAAAGCTG | ATTGGGAAGT | 900 |
| | CATATTACCT | GATATTAAAG | CTGCTTTTTC | TGATGCGAAT | AAGGTTTTAG | AATCTGTAAA | 960 |
| 30 | TGAACCTCAA | ATTGACAATC | ATTTTGGTGA | AATTAAAGCT | GAATTATTAA | CTTTTAAGGG | 1020 |
| | TATACCGTAT | CAAATTAAGC | TAACTTCTGC | TGACCAAGAA | TTAAGAGAAC | AATTACCACA | 1080 |
| 35 | AACATATGTT | GACCATATGA | CTCAAGCGCA | AACAGCACAT | GACAATATTG | TTTTTATGCG | 1140 |
| | TAAATGGCTA | GATTTAGGAA | ATCGCTATGG | AAATATTCAA | GAAGTAATGG | ATGGTGTCCT | 1200 |
| | AGAAGA AGTG | CTAGCTACCT | ATCCAGAATC | ACAGTTACCC | GTATTGGTAA | AACATGCTTT | 1260 |
| 40 | AGAAGAAAAT | CACGCAACTA | ATAATTATCA | TTTCTATCGA | CATGTCTCTT | TGGATGAATA | 1320 |
| | TCATGCAACT | GATAATTGGA | AGACTCGATT | ACGAATGTTA | AACCATTTTC | CAAAGCCGAC | 1380 |
| | TTTTGAAGAT | ATACCGCTGC | TTGATTTAGC | TTTATCTGAT | GAAAAAGTAC | CGGTTAGACG | 1440 |
| 45 | TCAAGCGATT | GTATTATTAG | GTATGATTGA | AAGTAAAGAA | ATTTTACCGT | ATTTATATAA | 1500 |
| | GGGGCTTCGT | GATAAAAGTC | CTGCTGTAAG | AAGAACAGCA | GGGGATTGCA | TAAGCGATTT | 1560 |
| _ | AGGGTATCCA | GAGGCACTAC | CAGAAATGGT | GCTACTATTA | GATGATCCAC | AGAAAATCGT | 1620 |
| 50 | TAGGTGGCGT | GCTGCTATGT | TTATCTTTGA | TGAAGGTAAT | GCAGAGCAGC | TTCCCGCACT | 1680 |
| | AAAAGCCCAT | ATTANTONCA | ATGCGTTTGA | ACTTABATTA | CARATTGAAR | TCCCCATATC | 1740 |

| | AATITAATTG | GAGGAATTAA | ATATGAATGC | ATATGATGCT | TATATGAAAG | AAATTGCGCA | 1860 |
|-----------|------------|------------|------------|----------------|-------------|------------|------|
| | ACAAATGCGT | GGCGAATTAA | CTCAAAATGG | TTTTACAAGT | TTAGAAACGA | GCGAACAGCt | 1920 |
| 5 | ATCGGAGTAT | ATGAACCAAG | TAAATGCTGA | TGACACTACT | TTTGTAGTTA | TTAACTCTAC | 1980 |
| | ATGCGGCTGT | GCAGCTGGAT | TAGCAAGACC | AGCTGCAGTA | GCAGTTGCAA | CACAAAATGA | 2040 |
| | ACATAGACCT | ACAAATACAG | TTACAGTTTT | TGCTGGGCAA | GATAAAGAAG | CAACTGCTAC | 2100 |
| 10 | AATGCGAGAA | TTCATTCAGC | AAGCACCATC | TAGTCCTTCG | TATGCTTTAT | TCAAAGGTCA | 2160 |
| | AGATTTAGTT | TATTTTATGC | CTAGAGAATT | TATCGAAGGT | AGAGATATTA | ATGACATTGC | 2220 |
| 15 | AATGGACTTA | AAGGATGCCT | TTGACGAAAA | TTGTAAATAG | TACACATAAA | TAAATATAAA | 2280 |
| | GGTTAACACA | TTTTATAATA | TTAAAAATGG | TGTCTGTCAT | TGAAAATAGA | GAATATAGTT | 2340 |
| | GTATTCTATT | TGTTAAATAA | AGTCCGTTTT | TACCAACTAT | ATTTTCTAGA | AATTTAACTG | 2400 |
| 20 | TTTTAATAGG | ACATCAAACA | TAATATTCaA | ATCATGTGTT | AACCTCTTTT | TTAAAATTTT | 2460 |
| | TTAGCATTAA | AGTTATAGAT | TTGGGTAAAC | AATTACCAAT | TGGAAACATA | TATCACGTTA | 2520 |
| | CGATGGGGTA | GGTACTTAAT | CAGCATTTTA | TAAATAAAGT | AACGGAATTC | ATGATATTAA | 2580 |
| 25 | TATCATATTC | CTAAAATGAG | TGATAACAAA | ATGCTACATA | AAGTTAAGTT | ATATCAAACT | 2640 |
| | AAATATACAT | ACTATAAATA | ATGAAAATGA | GGTGTTATCG | CATATGTTGA | ATTCATTTGA | 2700 |
| | TGCAGCATAT | CACAGTCTTT | GTGAAGAAGT | TTTAGAAATA | GGAAATACAC | GAAATGATCG | 2760 |
| 30 | CACAAATACA | GGTACGATTT | CGAAATTTGG | TCATCAACTT | CGCTTTGACT | TATCTAAAGG | 2820 |
| | ATTTCCACTA | TTAACGACAA | AGAAAGTTTC | TTTTAAATTA | GTAGCAACCG | AATTATTATG | 2880 |
| 35 | GTTCATTAAA | GGAGATACAA | ACATCCAATA | CTTATTAAAA | TATAATAATA | ATATATGGAA | 2940 |
| | CGAATGGGCT | TTTGAAAATT | ATATCAAATC | AGACGAGTAT | AAAGGTCCAG | ATATGACAGA | 3000 |
| | TTTCGGGCAT | CGTGCATTGA | GTGATCCTGA | ATTTAACGAA | CAATATAAAG | AACAAATGAA | 3060 |
| 40 | ACAATTTAAG | CAACGTATTC | TTGAAGATGA | TACATTTGCG | AAGCAATTCG | GGGATTTAGG | 3120 |
| | AAATGTTTAT | GGTAAACAAT | GGCGAGATTG | GGTTGATAAA | GATGGTAATC | ATTTTGATCA | 3180 |
| | ACTTAAAACA | GTAATTGAAC | AAATTAAGCA | TAATCCAGAT | TCAAGGCGAC | ACATCGTATC | 3240 |
| 45 | TGCATGGAAT | CCAACAGAAA | TTGATACAAT | GGCACTTCCG | CCTTGTCATA | CCATGTTCCA | 3300 |
| | GTTTTATGTC | CAAGATGGTA | AGTTAAGTTG | CCAGTTATAC | CAACGTAGCG | CAGATATCTT | 3360 |
| 50 | TTTAGGTGTG | CCATTTAATA | TCcGCagctA | CGCTTTATTG | ACACACCTTA | TTGCCAAAGA | 3420 |
| 50 | ATGTGGACTT | GAAGTGGGTG | AATTTGTGCA | TACATTTGGA | GATGCACATA | TTTATTCAAA | 3480 |
| | теататтаат | CCCATTCAAA | CACAATTACC | A CCTCA A A CC | TTCS STCCTC | CAACATTAAA | 3540 |

| | TGAATCACAT | CCAGCAATAA | AAGCTCCAAT | AGCAGTGTAG | TCATTGCATA | GTTAGCTAAC | 3660 |
|-----|------------|------------|------------|------------|------------|------------|------|
| | CATATAGACA | TCAAAATGAC | ATCATAGTAT | TTTCAAGTGC | AAAAAAGTAC | TTTTTTGTGT | 3720 |
| 5 | TAAACGTTTT | CATAAATTAT | GCAAAATCAT | TATTTCTATC | ACACTTTATG | ATAAAAATTG | 3780 |
| | TGTTAAATTA | AAGATAACTT | AGTAATAAAA | AATGAAATGA | TAGAAGAAGG | AGGATAATTA | 3840 |
| 0 | TGACTTTATC | CATTCTAGTt | GCACATGACT | TGCAACGAGT | AATTGGTTTt | GAAAATCAAT | 3900 |
| U | TACCTTGGCA | CCTACCAAAT | GATTTGAAGC | ATGTTAAAAA | ATTATCAACA | GGTCATACTT | 3960 |
| | TAGTAATGGG | TCGTAAGACA | TTTGAATCGA | TTGGTAAACC | ACTACCGAAT | CGTCGAAATG | 4020 |
| 5 | TTGTACTTAC | TTCAGATACA | AGTTTCAACG | TAGANGGCGT | TGATGTAATT | CACTCTATTG | 4080 |
| | AAGATATTTA | CCAACTACCG | GGCCATGTTT | TCATATTTGG | AGGGCAAACA | TTATTTGAAG | 4140 |
| | AAATGATTGA | TAAAGTGGAC | GACATGTATA | TTACTGTTAT | TGAAGGTAAA | TTCCGTGGTG | 4200 |
| 20 | ATACGTTCTT | TCCACCTTAT | mCATTkGAgr | CTGGGAAGTT | GCCTCTTCAG | TTGAAGGTAA | 4260 |
| | ACTAGATGAG | AAAAATACAA | TTCCACATAC | CTTTCTACAT | TTAATTCGTA | AAAAATAAGG | 4320 |
| | GGGAAAACGA | CCATGACAAA | ACAGATTATA | GTAACAGACT | CAACATCCGA | TTTATCTAAA | 4380 |
| ?5 | GAATACTTAG | AAGCAAACAA | CATTCATGTA | ATTCCTTTAA | GTTTAACTAT | TGAAGGAGCT | 4440 |
| | TCATACGTTG | ACCAAGTAGA | TATTACATCA | GAAGAATTTA | TTAATCATAT | TGAAAATGAT | 4500 |
| 20 | GAAGATGTAA | AGACAAGTCA | GCCAGCCATA | GGTGAATTTA | TATCTGCTTA | TGAAGAACTA | 4560 |
| 30 | GGAAAAGATG | GCTCTGAAAT | CATAAGTATT | CATCTTTCTT | CAGGATTAAG | TGGTACATAT | 4620 |
| | AACACTGCTT | ACCAAGCAAG | TCAAATGGTA | GATGCTAATG | TAACTGTTAT | TGATTCAAAA | 4680 |
| 35 | TCTATTTCTT | TTGGTTTAGG | GTATCAAATA | CAACACCTAG | TAGAGCTTGT | AAAAgAaGGT | 4740 |
| | GtCTCAACTT | CTGAAATAGT | TAAAAAGTTA | AATCATTTAA | GAGAAAACAT | TAAATTATTT | 4800 |
| | GTAGTTATAG | GGCAATTGAA | TCAATTAATT | AAAGGTGGCA | GAATTAGTAA | AACAAAAGGT | 4860 |
| 10 | TTGATTGGTA | ATCTTATGAA | AATTAAACCA | ATTGGTACAC | TAGATGATGG | TCGCTTAGAG | 4920 |
| | CTTGTGCmCA | ATGCGAGAAC | TCaAAATTCk | AGTATCCAAT | ACTTGAAAAA | GGAAATTGCT | 4980 |
| | GAATTTATAG | GAGATCATGA | AATCAAATCC | ATTGGTGTCG | CACATGCTAA | CGTCATTGAA | 5040 |
| 15 | TATGTTGATA | AATTGAAGAA | AGTTTTTAAT | GAAGCTTTTC | ATGTGAATAA | TTACGATATA | 5100 |
| | AATGTAACTA | CACCAGTTAT | TTCTGCACAT | ACTGGTCAAG | GTGCGATTGG | CCTCGTAGTC | 5160 |
| 50 | CTTAAGAAGT | AAATTTAATC | TTTTCAGTGT | TAATTACTTC | CATTTCAATC | CTTTATAGAC | 5220 |
| ,,, | TAAATTTATA | ATTAGATAGA | TAGAGGAGGT | AATTCATATG | ACAAAAGAAT | ATGCAACATT | 5280 |
| | AGCAGGAGGA | TGTTTCTGGT | GCATGGTTAA | ACCATTTACA | TCATATCCAG | GCATCAAGTC | 5340 |

| GAATCAAACC | GGCCATGTCG | AAGCAGTACA | AATTACGTTT | GATCCAGAGG | TTACTTCCTT | 5460 |
|------------|------------|------------|------------|------------|------------|------|
| TGAAAATATA | TTAGACATAT | ATTTCAAAAC | ATTTGACCCA | ACTGATGATC | AAGGGCAATT | 5520 |
| TTTCGATAGA | GGCGAAAGCT | ATCAACCAGT | CATTTTCTAT | CATGATGAAC | ATCAGAAAAA | 5580 |
| GGCTGCTGAG | TTTAAAAAGC | AACAATTAAA | TGAACAAGGT | ATTTTCAAGA | AACCAGTGAT | 5640 |
| TACACCTATT | AAACCATATA | AAAATTTCTA | TCCAGCTGAA | GACTACCATC | AAGATTATTA | 5700 |
| CAAAAAGAAC | CCGGTACATT | ATTACCAATA | TCAACGTGGT | TCAGGTAGAA | AAGCGTTTAT | 5760 |
| AGAATCACAT | TGGGGGAATC | AAAATGCTTA | AAAAAGATAA | AAGTGAACTA | ACAGATATAG | 5820 |
| AATATATTGT | TACACAAGAn | AACGGCACTG | AACCACCATT | TATGAATGAA | TATTGGAATC | 5880 |
| ATTTTGCTAA | AGGATTTATG | TAGATAAAnT | TCnGGTAAAC | CTTG | | 5924 |
| | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 131:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 9280 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

25

30

35

40

45

50

15 ,

20

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 131:

| GGCCGTTnAA | AATCTCCAAA | ATAnaaaaaC | CCATCTTGTT | CCAATGTTTT | AAAATCGCCa | 60 |
|---------------------|------------|-------------|------------|------------|------------|-----|
| TCCaACACTT | GaTCaATAGC | TTGCAACAAC | GTTGAACGTG | TTTTaCCAAA | AGCATCAAAC | 120 |
| GCTCCCACTA | AAATCAGTGC | TTCAAGTAAC | TTTCTCGTTT | TGACTCTCTT | CGGTATACGT | 180 |
| CTAGCAAAAT | CAAAGAAATC | TTTAAATTTG | CCGTTCTGAT | AACGTTCATC | AACAATCACT | 240 |
| TTCACACTTT | GATAACCAAC | ACCTTTAATT | GTACCAATTG | ÄTAAATAAAT | GCCTTCTTGG | 300 |
| GAAGGTTTAT | AAAACCAATG | ACTTTCGTTA | ATGTTCGGTG | GCAATATAGT | GATACCTTGT | 360 |
| TTTTTTGCTT | CTTCTATCAT | TTGAGCAGTT | TTCTTCTCAC | TTCCAATAAC | ATTACTTAAA | 420 |
| ATATTTGCGT | AAAAATAATT | TGGATAATGG | ACTTTTAAAA | AGCTCATAAT | GTATGCAATT | 480 |
| TTAGAATAGC | TGACAGCATG | TGCTCTAGGA | AAACCATAAT | CAGCAAATTT | CAGAATCAAA | 540 |
| TCAAATATTT | GCTTACTAAT | GTCTTCGTGA | TAACCATTTT | GCTTTGCACC | TTCTATAAAA | 600 |
| TGTTGACGCT | CACTTTCAAG | AACAGCTCTA | TTTTTTTTAC | TCATTGCTCT | TCTTAAAATA | 660 |
| TCCGCTTCAC | CATAACTGAA | GTTTGCAAAT | GTGCTCGCTA | TTTGCATAAT | TTGCTCTTGA | 720 |
| тааатаатаа | CACCGTAAGT | ATTTTTTAAT | ATAGGTTCTA | AATGCGGATG | TAAATATTGA | 780 |
| y Caracter Caracter | CATCATOTOT | ጥርጥጥርጥል ልጥር | таасттесаа | тттсттстат | TCCACCTCCT | 840 |

| | ACACTTCTTA | CACCGTCAGA | CTCTAATTGG | AATATGCCAG | TCGTATCTCC | TTGCGACAAC | 960 |
|----|------------|------------|------------|------------|------------|------------|------|
| | AATTCAAACA | CTTTTTGATC | ATCAAACGGA | ATCTTTTCGA | TATCAATATT | AATACCTAAA | 1020 |
| 5 | TCTTTTTTGA | CTTGTGTTAA | GATTTGATGA | ATAATCGATA | AGTTTCTCAA | CCCTAGAAAA | 1080 |
| | TCTATTTTTA | ATAACCCAAT | ACGTTCGGCT | TCAGTCATTG | TCCATTGCGT | TÄATAATCCT | 1140 |
| | GTATCCCCTT | TCGTTAAAGG | GGCATATTCA | TATAATGGAT | GGTCATTAAT | AATAATTCCT | 1200 |
| 0 | GCCGCATGTG | TAGATGTATG | TCTTGGTAAA | CCTTCTAACT | TTTTACAAAT | ACTGAACCAG | 1260 |
| | CGTTCATGTC | GATGGTTTCG | ATGTACAAAC | TCTTTAAAAT | CGTCAATTTG | ATATGCTTCA | 1320 |
| 5 | TCAAGTGTAA | TTCCTAATTT | ATGTGGGATT | AAACTTGAAA | TTTCATTTAA | TGTAACTTCA | 1380 |
| 5 | TCAAACCCCA | TAATTCTTCC | AACATCTCTA | GCAACTGCTC | TTGCAAGCAG | ATGACCGAAA | 1440 |
| | GTCACAATTC | CAGATACATG | TAGCTCGCCA | TATTTTTCTT | GGACGTACTG | AATGACCCTT | 1500 |
| 20 | TCTCGGCGTG | TATCTTCAAA | GTCAATATCA | ATATCAGGCA | TTGTTACACG | TTCTGGGTTT | 1560 |
| | AAAAAACGTT | CAAATAATAG | ATTGAATTTA | ATAGGATCAA | TCGTTGTAAT | TCCCAATAAA | 1620 |
| | TAACTGACCA | GTGAGCCAGC | TGAAGAACCA | CGACCAGGAC | CTACCATCAC | ATCATTCGTT | 1680 |
| 25 | TTCGCATAAT | GGATTAAATC | ACTTACTATT | AAGAAATAAT | CTTCAAAACC | CATATTAGTA | 1740 |
| | ATAACTTTAT | ACTCATATTT | CAATCGCTCT | AAATAGACGT | CATAATTAAG | TTCTAATTTT | 1800 |
| | TTCAATTGTG | TAACTAAGAC | ACGCCACAAA | TATTTTTAG | CTGATTCATC | ATTAGGTGTC | 1860 |
| 80 | TCATATTGAG | GAAGTAGAGA | TTGATGATAT | TTTAATTCTG | CATCACACTT | TTGAGCTATA | 1920 |
| | ACATCAACCT | GCGTTAAATA | TTCTTGGTTA | ATATCTAATT | GATTAATTTC | CTTTTCAGTT | 1980 |
| | AAAAAATGTG | CACCAAAATC | TTCTTGATCA | TGAATTAAGT | CTAATTITGT | ATTGTCTCTA | 2040 |
| 15 | ATAGCTGCTA | ATGCAGAAAT | CGTATCGGCA | TCTTGACGTG | TTTGGTAACA | AACATETTGA | 2100 |
| | ATCCAAACAT | GTTTTCTACC | TTGAATCGAA | ATACTAAGGT | GGTCCATATA | TGTGTCATTA | 2160 |
| | TGGGTTTCAA | ACACTTGTAC | AATATCACGA | TGTTGATCAC | CGACTTTTTT | AAAAATGATA | 2220 |
| .0 | ATCATATTGT | TAGAAAATCG | TTTTAATAAT | TCAAACGACA | CATGTTCTAA | TGCATTCATT | 2280 |
| | TTTATTTCCG | ATGATAGTTG | ATACAAATCT | TTTAATCCAT | CATTATTTTT | AGCTAGAACA | 2340 |
| 15 | ACTGTTTCGA | CTGTATTTAA | TCCATTTGTC | ACATATATTG | TCATACCAAA | AATCGGTTTA | 2400 |
| | ATGTTATTTG | CTATACATGC | ATCATAAAAT | TTAGGAAAAC | CATACAATAC | ATTGGTGTCA | 2460 |
| | GTTATGGCAA | GTGCATCAAC | ATTTTCAGAC | ACAGCAAGTC | TTACgGCATC | TTCTATTTTT | 2520 |
| 50 | AAGCTTGAAT | TTAACAAATC | ATAAGCCGTA | TGAATATTTA | AATATGCCAC | CATGATTGAA | 2580 |
| | TGGCCCCTTT | CTATTAGTTA | AGTTTTGTGC | GTAAAGCTGT | AGCAAGTTGC | TCAAATTCAT | 2640 |

| | CAATATCATT | AATAATCAAT | TGCCCTTTAG | AACGTAATCG | ACATCTGATT | TCATTACCTT | 2760 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | CATCGACTGC | AAATACCCAT | ATTTTCAAGC | CTTTGATGTC | AGCAATTGTA | TTAACAAACT | 2820 |
| 5 | GAGATGCTTC | ATTTGGCTGA | ATACCGAATT | GCTCCAATAC | ATCTTCAGTT | ATTTTAACTT | 2880 |
| | GGCAGAATCC | ATCATCCATA | AGTTCGAAAT | GTTGTAAAAC | ATAACCTTGA | AACGGCAACA | 2940 |
| | TTTTTGGGTC | CTTCTCCATC | ATTTATTTA | AAAGCGCATT | ATGATCAATA | TCATGCCCAA | 3000 |
| 10 | TTAACTTTCC | AGCAATTTCC | ATAGTATGTT | CTGAGGTATT | GTTAAAAAGG | AATCGCCCAG | 3060 |
| | TATCACCGAC | GATACCAAGA | TATAAAACGC | TCGCGATATC | TTTATTAACA | ATTGCTTCAT | 3120 |
| 15 | CATTAAAATG | TGAGATTAAA | TCGTAAATGA | TTTCACTTGT | AGATGACGCG | TTCGTATTAA | 3180 |
| | СТАААТТААТ | ATCACCATAC | TGATCAACTG | CAGGATGATG | ATCTATTTTA | ATAAGTTTAC | 3240 |
| | GACCTGTACT | ATAACGTTCA | TCGTCAATTC | GTGGAGCATT | GGCAGTATCA | CATACAATTA | 3300 |
| 20 | CAAGCGCATC | TTGÄTATGTT | TTATCATCAA | TGTTATCTAA | CTCTCCAATA | AAACTTAATG | 3360 |
| | ATGATTCCGC | TTCACCCACT | GCAAATACTT | GCTTTTGCGG | AAATTTCTGC | TGAATATAGT | 3420 |
| | ATTTTAAACC | AAGTTGTGAA | CCATATGCAT | CAGGATCTGG | TCTAACATGT | CTGTGTATAA | 3480 |
| 25 | TAATTGTATC | GTTGTCTTCG | ATACATTTCA | TAATTTCATT | CAAAGTACTA | ATCATTTTCA | 3540 |
| - | TACTCCCTTT | TTTAGAAAAG | TTGCTTAATT | TAAGCATTAG | TCTATATCAA | AATATCTAAA | 3600 |
| | TTATAAAAAT | TGTTACTACC | ATATTAAACT | ATTTGCCCGT | TTTAATTATT | TAGATATATA | 3660 |
| 30 | TATTTTCATA | CTATTTAGTT | CAGGGGCCCC | AACACAGAGA | AATTGGACCC | CTAATTTCTA | 3720 |
| | CAAACAATGC | aAGTTGGGGT | GGGGCCCCAA | CGTTTGTGCG | AAATCTATCT | TATGCCTATT | 3780 |
| | TTCTCTGCTA | AGTTCCTATA | CTTCGTCAAA | CATTTGGCAT | ATCACGAGAG | CGCTCGCTAC | 3840 |
| 35 | TTTGTCGTTT | TGACTATGCA | TGTTCACTTC | TATTTTGGCG | AAGTTTCTTC | CGACGTCTAG | 3900 |
| | TATGCCAAAG | CGCACTGTTA | TATGTGATTC | AATAGGTACT | GTTTTAATAT | ACACGATATT | 3960 |
| 40 | TAAGTTCTCT | ATCATGACAT | TACCTTTTTT | AAATTTACGC | ATTTCATATT | GTATTGTTTC | 4020 |
| | TTCTATAATA | CTTACAAATG | CCGCTTTACT | TACTGTTCCG | TAATGATTGA | TTAAAAGTGG | 4.080 |
| | TGAAACTTCT | ACTGTAATTC | CATCTTGATT | CATTGTTATA | TATTTGGCGA | TTTGATCGTT | 4140 |
| 45 | AATTGTTTCA | CCCATCTGAG | GCTGTCTTCC | TAAAAGTTGC | ATAGACTTTA | AAACATCTTG | 4200 |
| | TCTATTAATC | ACACCCACTG | TCTTTTTATT | ACTCGAAACG | ACAGGAATCA | ATTCAATACC | 4260 |
| | TTCCCAAATC | ATCATATGCG | CACAACTTGC | TACTGTACTC | ATAGCATTTA | CATAAATAGG | 4320 |
| 50 | ATTTCGCGTC | ATCACTTTAT | CTATTTCGTC | GTCGTCCTTT | GTATTAATCA | TCTCTCGACT | 4380 |
| | тсттасаата | ССТАСТВАТТ | татассастс | ATTGACTACC | GGAAATCTTG | TATGGCCAGT | 4440 |

| | ATCTAATGGC | GTCATTATAT | CTTGAACTAT | TAAGATATCT | TTTCGTATTT | TCTGATTAAA | 456 |
|----|------------|------------|------------|------------|------------|------------|--------|
| | AAGTGCTTTG | TTGATAATAT | TTGCAACTAG | GAATGTATCA | TAACTTGATG | ATAGAACAGG | 462 |
| 5 | TAAATCATGT | TCATTCGCAA | AATTAATAAC | TTTATTAGAT | GGCTTAAATC | CACCAGTAAT | 468 |
| | TAATATAGCC | GTACCTCTTT | TTAAAGCTTC | AATCTGCACA | TCTTCACGAT | TTCCGACAAT | 474 |
| | CAATAATGTC | TTTGGACCAA | TATACTTTAA | AATATCTTTG | AGTTCCATTG | CTCCAATTGC | 480 |
| 10 | AAATTTAGAT | ACCATCTTAG | TGATACCTTT | GTTGCCACCT | AACACTTGGC | CATCAATAAT | 486 |
| | ATTGACAATT | TCATTAAAAG | TTAAATGTTC | AATTTCATTA | CGATTACGTT | TTTCGATTCG | 4920 |
| | AACCGTACCA | ACACGATCTA | TCGTTGCGAC | CATGCCCATT | TTATCAGCAT | CTTTmATTGc | 4980 |
| 15 | ACGATATGCT | GTCCCytCaG | ATACGTTTAA | AAATTTAGCG | ATTTTACGCA | CCGAAATTTT | 5040 |
| | AGAGCCTATA | GATAACGATT | CAATATAATC | TAAAATTTGT | TCATGTTTTG | TCATTCTTTA | 5100 |
| 20 | CCTCTTCTTT | TCGAACAGTA | TTAACTACAT | TATAACTTTA | TTTTGGATAA | AAAGCATTGA | 5160 |
| | AGTGAAATGA | AATAATGATC | GTTLCACCTA | TTTTATTTTT | TGAAAATATA | CAACAAACAC | 5220 |
| | AAAGATCACA | AAATCTTTAA | TTTTAAATGG | AAAAATCCAT | TATTATTTAT | TAGAATGTAA | 5280 |
| ?5 | GTGAGGAGGG | ATGTACTAAT | GTATAAAAAT | ATATTACTTG | GTGTAGACAC | TCAGTTAAAA | . 5340 |
| | AATGAAAAAG | CACTAAAAGA | AGTGTCTAAA | TTAGCTGGCG | AAGGTACAGT | CGTAACAGTT | 5400 |
| | TTAAACGCAA | TCAGCGAACA | AGaTGCTCAA | GCATCAATTA | AAGCAGGTGT | TCATTTAAAC | 5460 |
| 30 | AAACTTACTG | AAGAACGAAG | CAAGCGATTG | GAAAAAACAC | GCAAAGCTTT | AGAAGATTAT | 5520 |
| | GGTATTGATT | ATGACCAAAT | AATTGTTCGT | GGTAATGCAA | AAGAAGAACT | ATTAAAACAT | 5580 |
| | GCTAATAGCG | GTAAATATGA | AATTGTTGTT | TTAAGTAACC | GTAAAGCAGA | AGACAAAAAG | 564 |
| 35 | AAATTTGTAC | TTGGAAGTGT | CAGCCACAAA | GTAGCAAAAC | GTGCGACTAT | CCCTGTATTA | 570 |
| | ATCGTTAAAT | AAAATTTTTA | TCCAGAATCA | CAAATAATCT | TTCAATCATG | ATGCAGTCTC | 576 |
| 10 | AAACGACTGA | GTÄAATACAA | GAAACGATTA | TGACTGTGGT | TCTGGATTTT | TTATATCGTA | 5820 |
| •• | GTAAATTTAT | AATCAATGTC | TAATTGTATA | AAACTAAAAT | TACGAGAGTA | GGTCAGAAAT | 588 |
| | GATAAAGAAC | CACTGATGTC | CCCCGTCCAC | GTCGTAACTG | AATCAGTAGA | ATATAAAAAC | 594 |
| 45 | ACCCACTAAA | AATATGCAGA | CGATAACTTC | CACATAGATT | AGCGAGGTGT | TTTTTAGTGT | 600 |
| | AAAATCTATA | TTCTATTTAA | AACTGAACAG | ATTCACCTGG | TTTTAAAATT | TGCACGTCCC | 606 |
| | CTACATTAAC | AGCATCTTTA | AATTGTTGTG | GATCTTGTTC | GATTAATGGG | AATGTATCAT | 612 |
| 50 | AATGAATCGG | TACAGAAATT | TTTGGTTTAA | TAAATTCATT | AATAGCATAA | CTTGCATCAT | 618 |
| | CAATACCCAT | CGTAAAATTA | TCTCCAATTG | GTACAAAACA | ТАСАТСААСТ | CCATCACCTT | 624 |

| | TTCAACTTCA | AACACGATAC | CCATTGGCAT | ACCTAAATAA | ACTGGGAATA | CCATTTTCAT | 6360 |
|----|------------|------------|------------|------------|------------|-------------|------|
| | GTGTAAAACT | TGAACTATGA | AATGCTTGAA | CAAATTTAAC | GCTTCCGAAA | TCAAaGTTTG | 6420 |
| 5 | CTTTACCACC | AaTATTCATA | CCATGAACAT | TTTCAACACC | GTGATATGAA | GAAAGATAGT | 6480 |
| | CAGCCATTTC | TGCACTTCCA | ATTACTGTTG | CTCCTGTTTT | CTTTGCTAGT | TCCACAACAT | 6540 |
| | CACCAAAATG | ATCAAAATGA | CCGTGCGTTA | AAACGATATA | GTCTACCTGC | ACTGTTTCAA | 6600 |
| 10 | TATTCAAATC | ACACTTAGGG | TTATTTGAAA | TAAACGGATC | TACGATAACC | TTTTTGTTGT | 6660 |
| | TCCCTTCTAA | ATAAATCGTT | GATTGACCAT | GAAATGATAA | CTTCATTTGA | GCATCCTCCT | 6720 |
| 15 | ATCAATTACT | ATATAAATTT | AGTACCCTTT | TGCCACTTAA | TTATAACAAA | TTCTCAAATT | 6780 |
| | TTAAAAATTG | AAAATCTAGT | TAATGTATTA | GCTCGATTTT | GAAATCTAAT | AATAATTGGC | 6840 |
| | ATAAAATGGA | AGTAATATTA | TGTTGAGGAG | TGTTTATAAA | ATGACAAAAA | TATCAAAAAT | 6900 |
| 20 | AATAGACGAA | TTGAACAATC | AACAAGCTGA | TGCAGCATGG | ATTACAACAC | CGTTGAATGT | 6960 |
| | ATATTATTTT | ACTGGATACC | GTAGCGAACC | CCATGAAAGA | TTATTTGCAT | TATTGATTAA | 7020 |
| | GAAAGATGGT | AAACAAGTAC | TATTTTGTCC | AAAAATGGAA | GTCGAAGAAG | TCAAAGCATC | 7080 |
| 25 | ACCTTTCACA | GGTGAAATCG | TTGGATATTT | AGACACTGAA | AACCCTTTTT | CACTITATCC | 7140 |
| , | TCAAACAATC | AATAAATTAC | TAATTGAAAG | CGAGCACTTA | ACAGTAGCAC | GCCAAAAACA | 7200 |
| | ATTAATCTCT | GGTTTCAATG | TCAATTCATT | CGGAGATGTT | GATTTAACAA | TCAAACAATT | 7260 |
| 30 | GAGAAATATT | AAATCCGAAG | ATGAAATTAG | CAAAATACGT | AAAGCTGCTG | AGTTAGCAGA | 7320 |
| | TAAGTGTATC | GAAATAGGTG | TTTCTTATTT | AAAAGAAGGT | GTGACTGAAT | GTGAAGTAGT | 7380 |
| | CAACCATATT | GAGCAAACTA | TCAAACAATA | TGGCGTCAAT | GAAATGAGTT | TTGATACGAT | 7440 |
| 35 | GGTTTTATTT | GGAGATCATG | CCGCATCACC | TCATGGCACA | CCAGGAGATC | GCAGATTAAA | 7500 |
| | AAGĆAATGAA | TATGTACTAT | TTGATTTAGG | TGTAATTTAT | GAGCATTATT | GTAGCGATAT | 7560 |
| 40 | GACACGTACT | ATTAAATTTG | GTGAACCTAG | CAAAGAAGCA | CAAGAAATTT | ATAATATTGT | 7620 |
| | ATTAGAAGCA | GAAACATCTG | CAATCCAAGC | AATTAAACCT | GGAATACCAT | TAAAAGATAT | 7680 |
| | CGATCATATC | GCTAGAAATA | TTATTTCAGA | AAAAGGTTAT | GGTGAATATT | TCCCTCATCG | 7740 |
| 45 | CTTAGGTCAT | GGCCTAGGAT | TACAAGAACA | TGAATATCAA | GATGTTTCAA | GTACTAATTC | 7800 |
| | TAATTTGTTA | GAAGCTGGCA | TGGTTATTAC | AATCGAACCA | GGTATTTATG | TACCTGGTGT | 7860 |
| | TGCAGGTGTA | AGAATTGAAG | ATGACATACT | TGTCACTAAT | GAAGGATATG | AAGTATTAAC | 7920 |
| 50 | ACATTACGAA | AAATAAGGAG | TGGGATAAAA | ATGAAAAGCT | TGTTACAAGC | GCATTCTCAT | 7980 |
| | TCAGTCAAAC | ACTGCCAATA | таасаттста | CCCCCTAACA | | TATCON ACTO | 0040 |

| | TGTAATGAAT | CAAATCAATA | TCATTCATGT | TCGATGATTT | CTTCGCATTG | TTTCTAGCTT | 8160 |
|----|------------|------------|------------|------------|------------|------------|------|
| | TAATTTATCA | TTATTTAATT | TTAATAACCA | AGGAGATGAT | AACGTCATTC | TTTAGTACGC | 8220 |
| 5 | TGTAATCCAT | TCCCTTTTCA | TCAAATTCAA | ATTATAATTG | TAATGCTTCT | TCTACAGATT | 8280 |
| | TATATTCCAT | TTCAAATGCC | TCTGCAACGC | CTTTATTGGT | TACGTGACCT | TTGTAAGTAT | 8340 |
| | TTAAACCTAA | TGATAATGGT | TGATTTGATT | TAAATGCTTC | TCTATACCCT | TTATTAGCTA | 8400 |
| 10 | GCATGAGCGC | ATAAGGTAGC | GTAGCATTAT | TTAAAGCTAA | CGTCGAAGTA | CGCGGTACTG | 8460 |
| | CACCTGGCAT | ATTTGCAACT | GCATAATGAA | CCACACCATG | CTTAATATAT | GTAGGATCAT | 8520 |
| • | CATGTGTCGT | AATTTTATCA | GTTGLTTCAA | AAATACCGCC | TTGATCAATA | GCAATGTCAA | 8580 |
| 15 | TAATAACTGA | CCCATTTTTC | ATTTGTTTAA | TCATGTCTTC | TGTTACAAGT | CTTGGCGCTT | 8640 |
| | TAGCACCTGG | AATTAAAACT | GCACCTATTA | CTAAATCACT | TTGTTTAACA | TACAACTCAA | 8700 |
| 20 | TATTCAACGG | ATTTGACATA | ATTGTATGTA | CACGTCCACC | GAATAAATCA | TCTAATTGTT | 8760 |
| 20 | GTAAACGCTT | TGGATTAACA | TCTAAAATCG | TAACATCŢGC | ACCTAGTCCT | AGTGCAATTT | 8820 |
| | TAGCTGCATT | TGTTCCTGCT | TGACCACCAC | CGATAATAGT | TACTTTACCC | TTAGGTACTC | 8880 |
| 25 | CTGGGACACC | ACCTAGTAGA | ATTCCCATAC | CACCATTAAG | TTTTTGTAGG | AACTCTGCGC | 8940 |
| | CAACTTGAGC | TGACATTCTT | CCTGCTACCT | CACTCATTGG | TGATAACAAT | GGTAAAGATC | 9000 |
| | GGTCTGGTAA | CTGCACAGTC | TCATATGCAA | TACTAATTAC | TTTTCTATCT | ATCAAAGCTT | 9060 |
| 30 | GTGTTAATTT | TTCTTCATTT | GCTAAATGAa | gatAaGTGAA | TAATACAAGC | CCTTCTTTAA | 9120 |
| | | | | • | AACCATATCC | • | 9180 |
| | | | | | GTAATCTACA | | 9240 |
| 35 | ATGATCCTGA | | | | | | 9280 |
| | | | | | | | 2200 |

(2) INFORMATION FOR SEQ ID NO: 132:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4669 base pairs
- (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 132:

| CTGATTAA | TC TCTTGTTGTC | GTGTATTTAC | TAATTGAATC | GTTGGTGTCT | GAACACGTCC | 6 |
|----------|---------------|------------|------------|------------|------------|----|
| CAGGGATA | GC TGTGCATCAT | ACTITGTTGT | TAGTGCACGC | GTTGCATTAA | TCCCAACAAT | 12 |
| CCAATCTG | CC TCACTTCTCG | CTAACGCTGC | ATAATACAAA | TOGTTATATT | GACGACCGTC | 18 |

55

| | ACGGATTGGC | TTTTTGTTAC | CAACTTTATO | CAAAATCAAT | CTTGCAACTA | GTTCACCTTC | 30 |
|----------|------------|---|---------------------|------------|---------------|------------|------|
| | TCGTCCaGCA | TCTGTTGCAA | TAATAATATC | TTTCACTTTA | TTATCTAAAA | TTAACGCTTT | 360 |
| 5 | TACTGTTTTA | AATTGTTTGC | TTGTTTTACC | AATAACAACA | GTTTTCATAT | ATTTAGGTAT | 420 |
| | AATTGGAAGG | TCTTCTAATC | GCCATTCCTT | TAAATTTTTA | TCGTATTGTT | CAGGTGTCGC | 480 |
| | ATTTGTCACT | AGATGACCTA | ACGCCCACGT | GACAATATAT | TGGTTATTTT | CAAAGTAACC | 540 |
| 10 | ATTACGCTTC | TGATTTATTT | GTAAAGCATC | AGCAATATCT | CTTGCGACTG | ATGGTTTTTC | 600 |
| | AGCTAATATT | AAAGATTTCA | TAAATTATCC | TTTCTCATAC | GTTCTTTTAT | TTCGAACGTG | 660 |
| 15 | CTTCATCTAT | TCCACTAATC | TTTGATTTAA | ATTCAATGAT | TGCAAATGAT | GTGTTAAATG | 720 |
| | TATTGTAAÇA | TGTTAATATC | ACTATTAACT | TTCATTŢCAG | TTGAAATACT | ATATAATAAA | 780 |
| | AGTAACAAAA | AGTACGGAGG | TAATGACATG | AGCATAGTTC | AGTTATATGA | TATTACACAA | 840 |
| 20 | ATAAAATCGT | TCATTGAACA | TTCGAATTAT | GAATCAGCAT | CATACTTATA | TAAACTTCCT | 900 |
| | CAACAGTACA | ATGAAATAGA | TGTATTAATA | ACCGATGCGA | TTGAATCACC | TGGTGTATTT | 960 |
| | TCGATTAAAG | AAAACGATTC | AATCAAAGCA | ATCATATTGT | CTTTTGCATA | CGATAAAAAT | 1020 |
| 25 | AAATTCAAAG | TCATAGGCCC | TTTCGTGGCT | GACAATTATG | TATTATCTGT | CGATACGTTT | 1080 |
| | GAAACGCTAT | TTAAAGCAAT | GACTTCGAAC | CAACCTGACG | ATGCCGTCTT | TAACTTTTCT | 1140 |
| | TTTGAAGAAG | GCATTCAACA | ATACAAACCA | TTAATGAAAG | TTATTCAAGC | AAGTTATAAC | 1200 |
| 30 | TTCACTGACT | ATTACATAGA | AGCCCGTACA | AGATTAGAAG | AAGATATGCA | CCAACCAAAT | 1260 |
| | ATCATTCCTT | ATCACAAAGG | GTTTTATCGT | GCTTTCAGCA | AATTACACAC | AACTACATTT | 1320 |
| | AAATATCAGG | CACAGTCACC | ACAAGATATC | ATTGATAGTT | TAGACGACCA | TCATCATTTG | 1380 |
| 35 | TTTTTATTTG | TTAGCGAAGG | TTTACTTAAA | GGTTATTTAT | ACCTTGAAAT | TGATTCACAA | 1440 |
| | CAGTCAATCG | CCGAGATTAA | ATACTTCAGT | TCTCATGTAG | ATTACCGTTT | GAAAGGTATC | 1500 |
| 40 | GCTTTCGAGT | TGCTTGCGTA | TGCATTGCAA | TATGCTTTTG | ATAATTTTGA | TATTAGAAAA | 1560 |
| <u> </u> | GTTTATTTTA | AAATTCGTAA | TAAAAATAAT | AAACTCATCG | AACGATTTAA | TGGTCTAGGT | 1620 |
| | TTCCATATCA | ACTATGAGTA | CATTAAATTC | AAATTCGAAT | CACGTAACGT | AAAAGATCAA | 1680 |
| 45 | ACAATCCCTG | AATAAAACAC | CAAGCAAATA | CCCTACAGTA | CATCATTAGC | ATGTATTGTG | 1740 |
| | GGTTTTTCTA | CTTTTTGTAA | ATATTGAAAA | TTATAAGTAG | TTGTTTTTTA | CTATTAGGGC | 1800 |
| | AGAATGCTTT | ACAATAACAT | GCAAGTGTCA | ATTAAGGGGA | GCACTTGCAT | AAATAGTATA | 1860 |
| 50 | GGAGAGTGAG | TAGTCTTGCA | ATTTCTTGAT | TTCTTAATCG | CACTTTTACC | TGCTTTATTC | 1920 |
| | TGGGGAAGTG | ערבידין אין די אין | ፐል አጥረጥረ ጥጥር | CTCCCCCCTC | CA COMMA CA A | 00111mmccm | |

| | TTCAATAATC | CTACTGTAAT | TATTGTCGGT | CTTATTTCTG | GTGCATTATG | GGCGTTTGGA | 2100 |
|----|---------------|------------|------------|------------|--------------|------------|------|
| | CAAGCGAATC | AGCTTAAATC | TATTAGTTTA | ATCGGTGTAT | CAAATACTAT | GCCAGTTTCT | 2160 |
| 5 | ACAGGTATGC | AATTAGTTGG | TACAACATTA | TTCAGCGTTA | TCTTTTTAGG | TGAATGGTCT | 2220 |
| | TCAATGACTC | AAATTATCTT | TGGTTTAATC | GCCATGATAT | TATTAGTTAC | TGGTGTAGCA | 2280 |
| | CTTACTTCAC | TTAAAGCTAA | AAATGAACGT | CAATCAGATA | ATCCTGAATT | TAAAAAAGCA | 2340 |
| 10 | ATGGGTATTT | TAATTGTATC | TACAGTTGGA | TATGTAGGTT | TCGTTGTACT | TGGTGACATC | 2400 |
| | TTTGGTGTTG | GTGGAACTGA | TGCATTGTTC | TTCCAATCTG | TCGGTATGGC | AATTGGTGGC | 2460 |
| 15 | TTTATCCTAT | CCATGAATCA | TAAAACATCA | CTTAAATCAA | CAGCACTTAA | TCTATTGcCA | 2520 |
| | GGTGTGATTT | GGGGAATTGG | TAACTTGTTC | ATGTTCTATT | CTCAACCAAA | AGTTGGTGTA | 2580 |
| | GCTACAAGTT | TCTCATTATC | ACAGTTACTT | GTTATCGTTT | CAACCTTAGG | CGGTATTTTC | 2640 |
| 20 | ATTTTAGGAG | AAAGAAAAGA | TCGTCGTCAG | ATGACGGGTA | TTTGGGCAGG | TATTATTATT | 2700 |
| | ATCGTGATAG | CTGCTATAAT | TCTAGGTAAT | TTGAAATAGA | AAGTTAAATA | CTCATGTAAC | 2760 |
| | GTAAAAATGT | AATCACTTCT | GAAAATAACC | ATTCACTTAT | AGAATGATTA | TTTAATTAAA | 2820 |
| 25 | TCGGGAATTT | TACGTTGAAT | GTTCCTCTAT | ATGTCCTAGG | AAATACGTGG | CTCTAAAAAC | 2880 |
| | AAAACGCAAT | AACACATCAT | GACATTAATC | ATGCGTTTTA | AGACTTTAAA | ATTAGCGATA | 2940 |
| | CTTTTAAAAT | CTTGATGATA | TTCATATATC | AAGTATGCGC | CATACATATG | AAGTGGATAG | 3000 |
| 30 | CTGCATAACG | CACTGCATTA | TCAACTTGAA | TGTATGAGTT | GAACAACTAT | GTCATAAATA | 3060 |
| | AAAGCCCCCT | TTTCACAATA | TACATTTACA | TATTGTGGTA | AAGGGGGCTC | TCATTTTCTA | 3120 |
| | CGAATACTAA | AATGGATTTT | ATTTTCAAAT | GTGTAAACTA | GACAAACACT | GCCTGATACA | 3180 |
| 35 | CGTACAAAAT | AATGATACTA | ATAATGATTG | TCAAATTGGT | CGTCATACCT | ATAAATGGCA | 3240 |
| | GTGTTCGATA | TTTAAACTGA | ATACCATAAG | AAATAATTGO | AACACCTACC | GGGAACATCC | 3300 |
| 40 | AAGTGACCAA | CAATGTCGTC | TTAATCATAT | CATCTGATAC | TGGTAACAAC | ACATATACTA | 3360 |
| | ACAATCCCGC | AACTAATGCT | AATCCATAAT | GCAAACATAA | ATATTTAATA | GTAGCAGGTA | 3420 |
| | TATACTTTCT | TTCCAGAGTA | AAATTCAACA | TGACACCTAG | CAAAATCATI | GATAACGGCA | 3480 |
| 45 | TATTTGCATG | GGAAAGTATG | CTAAAGAAAT | CGATTGCCAC | ATGTGGTAAA | TGGATGTGAC | 3540 |
| | TTATATTCAA | TATAAACATT | ACAATGTATO | TAACGAGTGO | CACTGATTG1 | AATAATTTCT | 3600 |
| | TACCTAAATA | TTTAAAATCG | AATTGATCA | TACCTTCACT | T AAAGTAGCTA | CCTACAAAGT | 3660 |
| 50 | AAGTAATTCC | AAACATCACA | AAGGCACCAG | CTATATCAG | CATAACAAA | TAAATAAGTC | 3720 |
| | CCCTTTTT & CC | · | TCNNTTNGT(| CATATGCAA | A CAATCCAATA | TTCATAGCAC | 3780 |

| | CAATCATTTT | CGCCACAATA | CCATATATAA | TCATTAAAAT | TGGTAAAATG | GAGAATGACA | 3900 |
|----|-------------|-------------|-------------|------------|------------|------------|------|
| | ATTTTAATTC | TGCACTGTTT | AAATTCACAA | TAACTAAAGA | TGGGAGTGTG | ACATTAAGAA | 3960 |
| 5 | CTAATGTAGC | AATGACTTGA | CTATCTGTTG | СТТТТАТААА | ATTAATGCGC | TTCAAAAAGT | 4020 |
| | AACCAAGCGC | AATTAATAAA | ATAATCATAG | TAAATTGTTC | TGTCACTGTT | ATCCCTTCTT | 4080 |
| | TCAATAATCT | TCATAATTTA | TAACTTTAAC | ATACTCCACA | GATATTTTAG | AAGTCTACTG | 4140 |
| 10 | TTTCATGCTA | TAATCTACAT | TAAATGCACT | TAATTATATT | TCAAAGGAGT | GTTATAGTAT | 4200 |
| | GTCTTTAGAA | AACCAACTAG | CCGAACTTAA | ATATGATTAT | GTTCGTCTTC | AAGGTGACAT | 4260 |
| 15 | AGAAAAACGG | GAATCTTTGA | ATTTAGATAC | TTCCGCACTT | GTTCGTCAAC | TTAAAGATAT | 4320 |
| , | TGAAAATGAA | ATTAGAAACG | TTCGTGCTCA | AATGCAAGAT | TAATAATCTA | TCATTCAAGC | 4380 |
| | AATAAATGCT | TTTTGTTACA | TAAATTTGAC | TAGCATTGCT | CTGAATACGT | TATATTGATG | 4440 |
| 20 | AATTGCTTCA | TTTTTCGCTC | AATTACATCT | AGAATCACAA | GATGTTGTCG | TGTTATGATT | 4500 |
| | TAGTGTTTCA | TTAACAACAT | ACACGCATAT | CTATCCCAAC | ACTGCTATTT | ATGTTTTCTA | 4560 |
| | CGCTGnTGTA | CTACATGAAC | CCTTTGAAAC | GGAGAGGAAG | TTATCATATG | CAATTTTAnC | 4620 |
| 25 | TGATTTTACT | AGCAATACTT | TAAChAATTG | nTAGTTTAAT | AGAATTTTA | • | 4669 |
| | (2) INFORMA | TION FOR SE | Q ID NO: 13 | 3: | | • | |

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2785 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

35

40

45

50

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 133:

| TITIGLACCCA | ICIGATACAA | TGCACCATGC | GGTTTAACAT | GATTAATTTT | AACTTGATGA | 60 |
|------------------|------------|------------|------------|------------|------------|-------|
| ATGCGACAAA | ACCCTTGTAA | TGCACCTAAT | TGATAAATCA | TCAAATTATA | AATCTCGTCG | 120 |
| -TTAGAGATAT- | CTATATTTCG | TCTGCCAAAG | CCTTTCAAAT | CAGGTAAACC | AGGATGTGCA | 180 |
| CCTACTGCAA | CATTATGTGC | TTTGGCAAGT | TTTACCGTTT | CATTCATTAC | ATTTTCATCA | 240 |
| CCAGCGTGAA | AACCACAAGC | AACATTCGCA | CTTGTAATTA | ACGGAATAAT | TTGATGATCA | 300 |
| CCACCAAAGG | AATAATTTCC | AAATGCTTCG | CCTAAATCAC | AATTCAAATC | AACTCGCATT | . 360 |
| ATAATTCCAC | CCCTTTAACA | ATTTGATGTT | TTTCTAAAAA | TTTAATATCA | ACATCTTTTG | 420 |
| CATCTCCATC | ACGATATAGT | GGATAATTTA | AAACTGCATA | TAAAAAATCG | GCAGTTGTAG | 480 |
| AAAATCCATC | TATCACCATT | TCATCTAAGG | TGACTTTCAA | CTTATCAATT | GCTGAAGCTC | 540 |

| | AACCGTGATA | TAGTAAAGAA | TCGACTCGCA | CATTAAAGCC | TTGAGGTAAA | TGTAACGCTG | 660 |
|-----|------------|------------|------------|------------|------------|------------|------|
| | TCACTTTACC | TGGTGTTGGT | TGAAATTTCT | TTTCaGGATT | TTCGGCATTT | ATTCTCGCTT | 720 |
| 5 | CTATCACATG | ACCATTAAAT | TGAATATCGC | TTTGTGAAAA | AGGTAAATGA | TTATGTTCCA | 780 |
| | ATAAATACAG | TTGTGCTGCA | ACCAAATCAC | GTTCTGCTCG | CATCTCTGTA | ACAGTATGTT | 840 |
| | CAACTTGTAT | TCGAGCATTC | ATTTCAATAA | AGTAATGTGC | GGTATCAGTT | ACTAAAAATT | 900 |
| 10 | CAATCGTACC | TGCACTTCTA | TAATTTGCTG | CACGTGCAAC | TTTAACAGCA | TCGTTACATA | 960 |
| | TTTGTTGTCG | TCTTTCTTCA | GTTAATGCTG | CACAAGGAGA | TTCTTCGATT | AATTTTTGAT | 1020 |
| 15 | TTTTACGTTG | TACAGAACAA | TCACGTTCCC | CTAAATGTAC | ATAATTATCC | TGCCCATCTC | 1080 |
| , - | CCATAACTTG | AACTTCAACA | TGTTTTGCAA | CAGGTATAAA | AGCCTCAACA | TAAACACGAT | 1140 |
| | CATCATCAAA | GTATTTTTT | CCTTCACTTT | TAGCTTCTTT | AAATGCCTTT | TCTAAATCTT | 1200 |
| 20 | CAGCTTTCTT | TACAATACGT | ATACCTTTAC | CACCACCGCC | ACTGGCAGCT | TTGATAACAA | 1260 |
| | CTGGATAACC | GATGTCTTTG | GCAAGATTCT | CAATTTCAGA | CACATGATTC | ACAGCACCAT | 1320 |
| | TTGATCCTGG | AATCACAGGA | ACACCTGCAT | GATGAACTGT | TIGICTIGCT | GTTATTTTAT | 1380 |
| 25 | CCCCCATCAT | TTCCATCGTT | TTTTTAGTAG | GCCCTATAAA | CGCTATGCCT | TGTTCCTCAA | 1440 |
| | CGGTTTGAGC | AAATTTTGTT | GATTCTGATA | AAAAGCCATA | TCCTGGGTGA | ATTGCATTAG | 1500 |
| | CACCAGTGAT | TTGTGCAGCA | GATATGATGC | GGTCAATATT | ТАААТААСТА | TCTAAAgCAT | 1560 |
| 30 | TArcwTCCCC | AATACATATA | GCTTGATCTG | CTAAATGTAC | ATGCAAGCTT | TGCTCGTCCC | 1620 |
| | CTTTTGCATA | AACTGCTACA | GTTTCAATCC | CATATTCTCT | GCAAGCTCTT | ATAATCCTTA | 1680 |
| | CAGCAATTTC | ACCTCTGTTC | GCAATTAAAC | AACGAAGCAT | TTACTTACCC | CCTTTACTTA | 1740 |
| 35 | ATACGTACCA | AAACTTGGTC | GTATTCAACA | TTTGTGCCAT | GATCAGCTAC | TATTTCAGTA | 1800 |
| | ATTICTCCAG | CAACATCTGT | TGTTACCTCG | TTTAATACTT | TCATCGCTTC | AACATATCCT | 1860 |
| 40 | ATAATATCTC | CCTTGTTAAC | TTTGTCACCG | ACATTCACAA | TTGGTTCAGT | TAATTCTTTA | 1920 |
| | CTATCTTGTA | AAAAGAATGT | ACCTATCATT | GGTGATTTAA | TGTCATGATA | ATCATTTGTC | 1980 |
| | GAAACATCGG | AGTTATCATT | CGCTTTTGAA | GCTGTCAAAT | CATTATTGTT | CATACTTTGA | 2040 |
| 45 | TTTGATTGAT | TACTGTGTGC | AGCCAAATGA | TTCGAGTCAG | TGAAGTCAAT | TTCTATTTCA | 2100 |
| | TCTTCAAAAT | TTTTATATTT | AAATTTCTTA | ACATCATTTT | CCTTCACTAA | TTTGATTATT | 2160 |
| | TGTTCGATTT | nTTCAATATT | CATTTTACAA | ATCCCCTTTT | AAAATTGTTG | CTAATTTTTT | 2220 |
| 50 | CGAAGTATGT | CGCAAGCTAG | ATGTATCAAA | AATTGGAGTC | TTTTGATGAC | TCTTAAGAAT | 2280 |
| | TTCATTAAAC | AGAGACATTT | GTTCCCGATT | CTTATCTACA | GCTTCTTGGA | ATGATATCCA | 2340 |

| TACAGTTGCA ATTTTGGTAT AACCACCTAT CGTTTGTTTA TCATTAAGCA GAATAATAG | G 2460 | | | | | |
|---|--------|--|--|--|--|--|
| TTGACCATCA TTTGGTACCT GAACACTACC AAGAGCAACC GGTTCAGAAA TGATATCTG | C 2520 | | | | | |
| TTGATTAALT GGTGCAACGC TGTCACCTTC CAAACGATAG CCCATACGGT CTGATTGTT | C 2580 | | | | | |
| AGTAATTAAA TATGGATGAT TTACAATTTT CGCTCTAGCC TCTTCAGAAA ATGCCTCGA | A 2640 | | | | | |
| TTGAGGTCCT TGAAGAATGT GTATAATATT ATTTTCTGGC AATAAATCGT CCTGTAAATC | G 2700 | | | | | |
| AATCGTCTTT CCAATGTTTT CTTTAAAGTC ATTATTTATT TTCACTGTTA TTACATCATC | 2760 | | | | | |
| AGCTAATAAC TTTCTACCTT TGAAT | 2785 | | | | | |
| (2) INFORMATION FOR SEQ ID NO: 134: | | | | | | |
| (i) SEQUENCE CURP CORPORATE | | | | | | |

. 5

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1010 base pairs
 - (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 134:

| AATGGAAACG | GTTGAAACAG | CAATTATTAC | TATTTCTATG | GGTGAAGGTA | TTTCAGAGAT | 60 |
|----------------|------------|------------|-------------|------------|------------|-----|
| ATTTAAATCA | ATGGGTGCCA | CACATATCAT | TAGTGGTGGA | CAAACGATGA | ATCCTTCTAC | 120 |
| AGAAGATATC | GTTAAAGTCA | TTGAACAATC | AAAATGTAAA | CGTGCAATTA | TTTTACCGAA | 180 |
| TAATAAAAAT | ATCTTAATGG | CAAGTGAACA | AGCAGCGAGT | ATTGTTGATG | CAGAAGCTGT | 240 |
| TGTTATTCCA | ACGAAATCTA | TTCCTCAAGG | TATAAGCGCA | CTATTCCAAT | ATGATGTGGA | 300 |
| CGCAACACTT | GAAGaAAATA | AAGCGCAAAT | GGCTGATTCA | GTAAATAACG | TTAAATCTGG | 360 |
| TTCATTAACG | TACGCTGTTC | GTGATACGAA | AATTGATGGC | GTTGAGATTA | AAAAAGACGC | 420 |
| GTTTÄTGGGC | TTGATTGAAG | ATAAGATTGT | AAGCAGCCAA | AGTGATCAAT | TAACAACGGT | 480 |
| TACTGAGTTG | TTAAATGAGA | TGTTAGCAGA | AGATAGTGAA | ATATTGACTG | TGATTATTGG | 540 |
| TCAAGATGCA | GAGCAAGCAG | TTACAGATAA | -CATGATAAAC | TGGATCGAAG | AGCAATATCC | 600 |
| AGATGTAGAA | GTGGAAGTTC | ATGAAGGTGG | ACAACCAATT | TATCAATATT | TCTTTTCAGT | 660 |
| AGAATAAAAA | TTTAAAATAA | AAAACTACCA | ATGATAAATC | ATCAGTTGGT | AGTTTTTTAT | 720 |
| TTTGCTATTT | TAGTGATATT | GCGGGTTAAA | AGTATCGTTC | TCGAGTTGCT | AACAATGTCA | 780 |
| TGTTCAACTT | AGTCATGATA | AAATAAATAA | CATACTAAAT | GATACGTAAA | ATCAAATAAA | 840 |
| ACATAGGTGA | TTTATTTTGG | CTAAAGTAAA | CTTAATAGAA | AGTCCATATT | CTCTTTTACA | 900 |
| ATTAAAAGGT | ATAGGTCCTA | AGAAAATAGA | AGTATTGCAA | CAACTAAATA | TTCATACAGT | 960 |

(2) INFORMATION FOR SEQ ID NO: 135:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1540 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

10

15

20

25

30

35

40

45

50

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 135:

| 60 | AGCCATTTTG | AAGCAATACA | AAAGACAAAA | AAAAGAAAGA | CATGAACAAC | TGTAGTTGAA |
|------|--------------------------|------------|------------|------------|------------|---------------------------|
| 120 | CTTCAAGTCT | CTCCCGGCAC | TGTACTATTA | TACTTATAGT | AGTTTTATAA | GATTGTCATG |
| 180 | TGTGGGTAAC | GCAGTTATTA | TTWAGCTTTT | TACTAGCTAT | GCTAAGGCAG | GCCGATAATG |
| 240 | TATTACTTTT | GGCTTAATGA | TTTAATTATT | TGTCAGCAAC | TCATATCCGG | GGAAGCTGTA |
| 300 | GTGGCAGTGC | AATCCGAAAA | GAAGCTAGGT | ATTTAGGGGA | CCTGTTCAAA | AGGATTTAGC |
| 360 | CGTTTAGTGG | TTATCATTAG | TAATCATGCA | TTCTAGGAAC | GGAAGTGACC | TATTTTAGCT |
| 420 | CTGCTATGCA | TTTTTGGCTG | AGCTGCATTA | CTCTCGTTGC | TCAGCTGTAG | ATTTGCAACT |
| 480 | GTAATAAAAC | TCAATTGTTG | TTTAGTGTTA | GACTAGCTCT | TTGCATAAAA | AGAAACGAAT |
| 540 | TCGTTCCTTC | CTTGCATTTT | TTCAATTGTA | CAATTATCGT | GTTATTGGAG | TAGAAATATA |
| 600 | CGGCATTTAA | GGTATGATTG | AATCTTGCTG | CAGTTGTACC | AGAGCAGGGG | TGCAACAGCT |
| 660 | AAGCTGTGTC | ACTTCAGTAC | ATTAATAATT | TAGCGTCTTT | GATAGCAAGT | AGTTTCCAAA |
| 720 | TTAATTTTAT | ATCGTAGCGA | AGCACAAAAT | AAACGGCGGC | ATTGGTATCA | aatttggaat |
| 780 | CAGCGCCTTG | TTCTTATATG | GGGCGAGTGG | ATGTTTCATG | TTAGGATTTG | AAACCATCAA |
| 840 | CTCCAGAAAT | AAAGTGATGC | CATCATGATT | CTTTATATTT | ATGTCCGTAG | GTCCATAGTT |
| 900 | TTGGCCCCGT | TTGCATAAAC | AAAAGAAGAA | AAGATTTAAT | GAAGGTGGTA | ТААТ <mark>А</mark> СААТА |
| 960 | TTTGGTCAAC | TTATTACTGT | ATCGATGTTA | TAATTGTTAT | GAATGGCGTT | TAGCCCACGT |
| 1020 | TAGGTGTTAT | ATTATTGCTT | ATCCATTACT | TTGACTCTGC | TTACATCCGA | TGAAAAAGTA |
| 1080 | TACCATGGGG | GAAAATAAAA | GAAACATGTT | TCATGACATG | AAAATTGGTG | GTTAATGCCG |
| 1140 | AAACAGGTGC | GTTCTTTTGA | ACTAGGTAAC | TAGGTATTTC | GTGTTTGGTG | AACAATTATC |
| 1200 | TACCTATTAT | TTAAAACATT | TGTTTTAGGT | AAACTTTTGG | TTAAGTGATC | AGCTCAATGG |
| 1260 | TTGCGAGTGC | CATTTGGGCT | TATATTGATT | CGCTTTTTAA | GCACTTATCA | CGCGACAATT |
| 1320 | CGTTACACTT | CTAACCTCTA | TTTTATTTCG | TAATACCTGT | TCATCAGCGT | AACAAGTTTA |
| 1380 | Հար Հարաստանի | GTTATTAGTT | TCAACAATTT | TTGTTTTAAT | TCTATAGGAT | AGGAGACCAG |

· 55

| | AGAIIICIIG AAGGCAGGTA TACCATTGAC AATTGTAGGG AATALCLAGT GATAGTTTTT | 1500 |
|-----------|---|------|
| | AGCATGACTT ATTGGAAATG GGTAAGGTTG CNTTAATTAA | 1540 |
| 5 | (2) INFORMATION FOR SEQ ID NO: 136: | |
| | (i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 11823 base pairs(B) TYPE: nucleic acid | |
| 10 | (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 15 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 136: | |
| 75 | ACTTCTCACA ATAAGAAATA TGAAATTGTT ATGTGTTAGT TGAGATTCAG TGATGAATTA | 60 |
| | CTTTTATCAT TTAAAATGTT GTTATCATTG TCATGCGTTA CCAAATCGCT TACGTATACA | 120 |
| 20 | CGATTCCCAA TCTTAACATA GACGATTTGT ATATCAGAAT TITCTGATTA CTAACAGTTT | 180 |
| | ACCTAAGTTT AAATATCTGT TCAATGATTT TCAGTTATTT TTAAAAGAAA AATCGTAATG | 240 |
| | CTGCCATGAT AACAATCCCA CTAATAATTG TAATAGTTAA ALACGCGTGA TTATAGATAA | 300. |
| 25 | AATAACCGTC GGAATGAGCG CGATAATGTA AGGGATGTTT AATGTATACC CCTCACCATG | 360 |
| | AGGCGTCTGT TGAATAATGC TGTCAATGAC AAGTGCCGTA AATAGTGTGA TTGGGATAAA | 420 |
| | TGATAGCCAT CGAACCACGA CATCAGGCAA TTGCACTTTT GAAATCATGA TAAAAGGTAT | 480 |
| 30 . | AATTCGAATT AATAGCGTTA CGATACCACA CAATAAAATA AGTATTAACA TGTTCATATG | 540 |
| | AGTTATCATT GTTCCATCAT CACTCCTAAC GCTGCTGAAA TTGTGGCTGC AATTAATATT | 600 |
| | GCTAGATATG AAGGCATAAA CATACTTAGC GATAACATCA TTACTATGAC GGCAATAATG | 660 |
| 35 | AGTACTATGT AAATTCTTAA TCGCGATTTA GTAATTGATT CAAATTGCGC AATGGCCAAA | 720 |
| | AAGATAAACA TAGCCGTGAT AGCAAAATCT AACCCTAGCG TTTGCGGATT TGAGATATAT | 780 |
| 40 | TCGCCAAATA AAGCCCCAGC TACACATGAA ATTGCCCCAAA ATAAATATGC TGTGATGTTA | 840 |
| | AGACCATGCA TCCAACGATC ATTGATAGCT TCTCCTTTTA AATAAGGTGT AATGGCGACG | 900 |
| | CCAAACGTTT CGTCAGTTAC TAATGAACCT AATCCAACAC GGTTCCAAAA CCCATATGTC | 960 |
| 45 | TTGAAGTTTG GTGCAAGCGA CATACTTAAA AGGAACATTC TTGAATTTAC GATAAATACA | 1020 |
| | GTTAGTACAA TCGCTGATAT AGGTGTACCT GCTATAAACA ACGCGCACAT AATAAATTGC | 1080 |
| | GCAGCACCGG CATATATAAC AAGACATAAC AAGACAATTT CTAAAATACT AAAGTTTTGA | 1140 |
| 50 · | GACGAAGCCA CAATACCAAA TGAAATACCA ACACCGGCAT AACCCAATAA TGTTGGGATA | 1200 |
| | CACTCTTGCA CGCCTTGTCT AAAACTTAAA TGTGTTGTCA TCTCAATTAC CTCCTTTGCC | 1260 |

| | TAAGCAATAA | CATTAGACAT | CAGTTTGTCT | GAGGTTAGAC | ATTCCGGAGT | CTTTAGTCAG | 1380 |
|----|------------|-------------|--------------------|------------|--|------------|------|
| | CTTCATATTA | ACTITITATI. | TTTGAGAATT | TTCAATTTTT | TATTTAAGAC | TACCTCCATA | 1440 |
| 5 | TTTTCTATGG | aTTTGTAGTT | GTTTTTAAGT | ATCAATTTTA | TAAATTTTTA | TATCTGATGA | 1500 |
| | TGAGTCTGGG | aTATTGaTTC | ATGTACCACT | CCCTTaTaAT | CATCCCCTCC | CCCTaCCCTA | 1560 |
| | CTCCATCGAT | ATAACTCATA | CTACATATCA | ACGAAATCAG | TATTTTATCG | CTTCCTTTCC | 1620 |
| 0 | TATATTAGTG | ATGCTCAAAC | TTGTTACGTT | TTAGATTGTT | TTAGTTCATC | ATAATTATCC | 1680 |
| | CGTATTGTTG | CTATAATGAA | ATGCGTTCAC | CCCATTAAAC | CACAAACTTA | ATTTATTGTT | 1740 |
| _ | GTTATGTGCA | TTGGCTCACT | ATTATATTTT | TACAGCACAA | AAAAAGTGGC | GACAGTTCGT | 1800 |
| 5 | CACCACTTTT | TAAAATATTA | TTTAAAGTAT | CTTGCCCTTG | CTTTAAGTAT | ACGTAGATAT | 1860 |
| | ATACTTTTTA | AAGCTTGTAG | CTAAAGCCTT | TATTTAACTG | GTTTTGAAAT | TTGTGTTTTA | 1920 |
| 20 | CCACCCATAA | ATGGTACTAA | TGCTTCTGGA | ATTGTTACTG | TTCCATCTTC | ATTTTGGTAA | 1980 |
| | TTTTCAACAA | TAGCAGCAAA | TGTACGTCCA | ACTGCTAAAC | CACTACCATT | TAATGTATGT | 2040 |
| | GCTAATTCTG | GTTTAGCTGC | TTTGTCACGC | TTGAAGCGGA | TGTTAGCACG | ACGCGCTTGG | 2100 |
| 25 | AAATCCGTAC | AGTTTGAGCA | TGAACTAATT | TCTTTATAAT | CATTGTAGCT | TGGTAACCAA | 2160 |
| | ACTTCTAAAT | CATATGTTTT | GCTTGCACTA | AATCCAATAT | CACCTGTACA | TAAAATAACA | 2220 |
| | CGACGGTATG | GTAAACCTAA | CTCTTCTAGA | ATTGCTTCTG | CGTTTGTTGT | CATTTCTTCT | 2280 |
| 30 | AAAGCATTCC | ATGAATCTTC | AGGTTGTTC A | AAACGTACCA | TTTCCACTTT | ATCGAATTGA | 2340 |
| | TGTAAACGAA | TTAATCCTCT | TGTATCTCTA | CCTGCTGATC | CTGCTTCACT | ACGGAAACAT | 2400 |
| | GCAGATTGAC | CAGTGAATTT | TTCAGGAAGT | ACACCTGGTT | GAATAATTTC | ATTACGGTAG | 2460 |
| 35 | AAATTCGTTA | ATGGTACTTC | AGCAGTTGGA | ATTGTATATA | ATCCTTCTTT | TTCTACTTTA | 2520 |
| | AATAAATCTT | CTTCAAATTT | AGGTAATTGA | CCTGTACCAT | ACATTGTATC | TGCGTTCACA | 2580 |
| | AGCTGTGGTA | CCATCATTTC | TGTATAACCA | TGTTGTGTTG | TATGTTTTGT | AATCATATAG | 2640 |
| 40 | TTCATTAAAG | CACGCTCTAA | TTGCGCACCT | TCATTTGTTA | AATATACAAA | ACGCGCACCT | 2700 |
| | GAAACTTTTG | CTGCACGATC | AAAATCAGCC | ATTTTCAATT | CTTCTACAAT | ATCCCAATGT | 276 |
| 45 | GCTTTGGGTT | CAAATGAAAA | CTCaCGTGGT | GTACCCCACT | TTTTAACTTC | AACGTTATCT | 282 |
| | TCATCAGATT | CACCTTGAGG | TACATCATCA | CTTATTAAAT | TTGGAATACG | ACAAAGGATA | 288 |
| | CCTGTCATTT | TATTATCAAT | TTCATTTAAT | TGACTATCTT | TTTCTTTAAT | ATCGTCACCT | 294 |
| 50 | AATGTGCGCA | TTTCAGCAAT | CACATCATCA | GCATTTTCTT | TATTACGTTT | TTTTAATGCG | 300 |
| | > mmommoco | mm | 1 001 00m00m | | 00000000000000000000000000000000000000 | | 306 |

| | TCAATTTTGC | TCTTAACTGT | GTCAGGCTCA | TTTCTGAATA | ATCTAATGTC | TAACATTAAC | 3180 |
|----|------------|--------------|------------|------------|------------|------------|------|
| | CTTCATCCTT | TCCCAAATAA | TTATCATTTA | TTATGGAATG | ACGTACGTCT | TTATTTTTTA | 3240 |
| 5 | GAAAATAAAA | AAAGACCACA | TCCCTACAAG | GGACGTGGTC | TACGCGTTGC | CACCCTATTT | 3300 |
| | AACAATTTAA | GTTATAAAGA | TACACTAAAC | CTAAATTGCA | CTTCACTAAA | ATAACGGTTA | 3360 |
| | TCACCGATTG | TTCTTTTAAA | TTAAGTAGGT | AGATTCATAT | ATATGTTGAT | TCTTGTTCAC | 3420 |
| 10 | ACTAACCACA | AGCTCTCTGA | TATCGAACAC | TATATATTAC | TTGTCCTACG | AACAATGTCT | 3480 |
| | TATTAAGTTA | TTTTTAATAT | AGCAAACTAT | ATTTGCTTTT | TCAAGTAACG | ATTTCAAACA | 3540 |
| 15 | TCACTCATGT | CGATTTAGTG | ACATGCAGTC | GTTTGATAAA | TTGATTGCTT | TAAATACTGT | 3600 |
| | GCAACCGCTT | CAATATCTTT | ATGAAATTGA | CGATCATGTG | TAATGGATGG | CACGATACTT | 3660 |
| | CGAAACTCAT | CATACTTGCG | ACGTGTTTTT | GGTGATAATĆ | CTTCAACACC | TTTTAACTCT | 3720 |
| 20 | GCTGCTTGTA | ATGCAATAAC | ACATTCGATT | GCCAGCACAC | GTCTTGCATT | TTCAATAATT | 3780 |
| | TGATAACCAT | GTCTAGCAGC | TGTAGTTCCC | ATAGATACGT | GATCTTCTTG | GTTCGCAGAT | 3840 |
| | GAAGTGATAG | AATCAACACT | CGCTGGATGC | GCTAAAGTTT | TATTTTCAGA | AACGAGACTT | 3900 |
| 25 | GCAGCAGCAT | ATTGCATAAT | CATCGCGCCA | CTTTGCAATC | CTGGCTCTGG | ACTAAGAAAT | 3960 |
| | GCTGGTAAAT | CACCATTTAA | TTGAGGATTT | ACTAGTCGCT | CTAGACGACG | TTCCGATACG | 4020 |
| | TTTGCTAATT | CACTTACACC | TAATTTAAGA | TGATCTAATG | CAAAAGCAAT | AGGTTGTCCA | 4080 |
| 30 | TGGAAGTTAC | CACCTGAAAT | AACAAACGTT | TCATTTGCTT | CCTCAAATAT | AAGTGGATTA | 4140 |
| | TCATTAGCCG | CATTCATTTC | AAATTCTAAT | TGCTGTTTAA | CATAATTGAA | TACTTGAAAA | 4200 |
| | CTCGCGCCAT | GGATTTGTGG | TATACAACGC | AACGTATATG | CATCTTGTAC | ACGTATTTCT | 4260 |
| 35 | GATTGTCGCG | TCGTTAATGT | TGATCCTTCT | AACCAATCAC | GCATACGCGC | TGCCACATTA | 4320 |
| | ATCTGTTCTT | GAAAATTACG | AACTGCGTGC | ACATCATGTC | GATATGCATC | TATAATGCCA | 4380 |
| 40 | TTAAGAGACT | GATGCGTTAA | TGCAGCAATC | CATTCAGATT | GGTAACCTAA | ATCTTCTGCT | 4440 |
| | TCTATATAAC | TAATGACACC | TTGAGCTGTC | ATAGCTTGCG | TACCATTAAT | CAATGCTAAA | 4500 |
| | CCTTCTTTAG | CCTGAAGGTT | CAAAGGTTGT | CTATTTAATT | CTCTTAATAC | ATCGTCACTA | 4560 |
| 45 | TCCTTTTCTT | CCCCTCTGTA | CAATACTTTC | CCTTCACCAA | TTAATGCTAA | TGCTAAATGT | 4620 |
| | GATAATGGCG | CTAAATCTCC | TGATGCACCG | AGAGAGCCTT | GCTGTGGGAT | TATCGGTATA | 4680 |
| | ATACGTTCAT | TTATAAAAAA | TTGTAATTGT | CTCACTAATT | CTAAAGTGGC | ACCTGAATGA | 4740 |
| 50 | CCTTTTAATA | ATGTATTCAA | TCGTAAAATC | ATCATGACTA | ATGCTACTTC | TITTGAAAAT | 4800 |
| | | cmcc1 c1 ccc | | | ~~~~ | 1001010000 | |

| | TCCTCATTTT | CAATAATACG | TTCAACTACC | GCTCTACTTT | TTTTGACACG | TTCTAACGCA | 4980 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | TCATCAATAA | TTTCAATCTT | TGATTGTTGT | TGTAAAAATG | ATTTAATATC | CTCAATTGTT | 5040 |
| 5 | AGTGTTTCAC | CATCTAAATA | TAAAGTCATA | TATGTTACCC | CCTTGTTTAT | ATTAAGTAAC | 5100 |
| | CCATCCTTCT | TGAAGTATAC | GTTTTCATTT | TTATTGAAAC | AATGGTTTTA | CGTACATTTA | 5160 |
| 10 | TAACCTATTA | TCAGAGCACT | ATTGTAGTGC | GTTAAAGGAT | ATTAAGATTG | TTGTAAGCAT | 5220 |
| 10 | ATTTAATAAT | TTATCTATTG | ACGAATTGCA | TATACAGGTA | TAGTATTTTC | TATTGTATTT | 5280 |
| | AACGACAAAT | AATAATGAAT | TCAGAAATTT | ATAATACATT | TTGTTAAAAG | TTACTATATA | 5340 |
| 15 | TTTTTAAAAT | TGAATAAATT | CGGAAAAGGC | TTTTACATGG | GAGGTTATAT | CACTATGGAA | 5400 |
| | ACGTTAAATT | CTATTAACAT | TCCTAAGCGT | AAAGAAGATT | CACATAAAGG | TGATTATGGC | 5460 |
| | AAAATTTTAT | TAATTGGTGG | ATCTGCTAAC | TTAGGTGGTG | CCATTATGTT | AGCGGCTCGT | 5520 |
| 20 | GCATGTGTAT | TTAGCGGTAG | TGGTTTAATC | ACTGTAGCTA | CACATCCAAC | AAATCATTCA | 5580 |
| | GCATTACATT | CTCGTTGCCC | AGAAGCGATG | GTTATTGATA | TTAATGATAC | GAAAATGTTG | 5640 |
| | ACGAAAATGA | TTGAAATGAC | TGACAGTATA | CTAATTGGTC | CAGGTCTTGG | CGTTGATTTC | 5700 |
| 25 | AAAGGAAATA | ATGCCATTAC | ATTCCTACTA | CAAAATATAC | AACCGCATCA | AAATTTAATC | 5760 |
| | GTAGACGGCG | ATGCGATTAC | AATCTTTAGT | AAACTGAAAC | CGCAATTACC | TACATGTCGT | 5820 |
| | GTGATCTTTA | CACCACACCT | CAAAGAATGG | GAACGATTAA | GTGGTATTCC | TATTGAGGAA | 5880 |
| 30 | CAGACATATG | AGCGTAATCG | TGAAGCAGTT | GATCGTTTAG | GTGCAACTGT | TGTACTTAAA | 5940 |
| | AAACATGGTA | CTGAAATTTT | CTTTAAAGAT | GAAGACTTTA | AATTGACAAT | CGGTAGCCCA | 6000 |
| | GCAATGGCGA | CTGGTGGTAT | GGGCGATACA | CTTGCTGGTA | TGATTACAAG | CTTTGTCGGT | 6060 |
| 35 | CAATTTGATA | ACTTAAAAGA | AGCGGTTATG | AGTGCCACAT | ATACACATAG | TTTTATTGGC | 6120 |
| | GAAAACCTTG | CAAAAGATAT | GTATGTGGTG | CCACCATCAA | GACTTATCAA | TGAAATACCT | 6180 |
| | TACGCAATGA | AACAATTAGA | AAGTTAGTCA | TTACTAATCA | TTGAATATAG | TAAAGCATTA | 6240 |
| 10 | CTTTCTAGCA | TAAAAATAAG | ACTCCCCTAC | ATATAGGGAA | GTCTTATTTT | TTATTATTCT | 6300 |
| | TCATCTGATG | ATTGTTGTAT | ATCTTCTTCA | ACACGATCCA | TGAAATCTTG | TCTTACTTCA | 6360 |
| 15 | ATACGTCCAT | CTTCATCATT | TTCTTCTGAA | TCAATCACTT | CAGTATGAAT | TGCATTTCCT | 6420 |
| •• | GGTGTTTCAT | CATTTACAAC | CGCTTCACGT | TGTTGTTCAG | TACCATCTTC | AGATACAGTT | 6480 |
| | GAAGTAGATT | GCTCATCTTC | ATTCGTTTCA | TCTTCTGCAT | CTTCTTTTAC | TTTAGCAACC | 6540 |
| 5 <i>0</i> | GTTGAAACAA | ATTGATCATC | ACCTAAGCGA | ATTAAGCGAA | CACCTTGTGC | TGCACGACCA | 6600 |
| | TTTTGAGAAA | TATCTGCAAC | ATCTAGTCGA | ATAATGACAC | CTGCATTAGT | AACAATCATT | 6660 |

| | GTAGCTGTTT TAATACCTTT ACCACCACGA TTTGATAAGC GATAGTCATT AACTGGCGTA | 6780 |
|------------|---|------|
| 5 | CGTTTACCAT AACCATTTTC AGTAACTACT AATACTTCAT CAACACTGTT TGCATGAGCT | 6840 |
| J | ACATCAAGCC CTACAACTTC GTCACCTTCA CGAAGTGTAA TACCTTTCAC ACCCGTTGCT | 6900 |
| | GTACGGCCTA AAGGACGTAA TGTTGATTCA GGGAATCGAA TTAATGATGC ATGTGATGTA | 6960 |
| 10 | CCAATCAAGA TATCTTCTTG ACCACTTGTT AAGCGAACTG CAATTAACTC ATCATCTTCT | 7020 |
| | CTGAACGAAA TCGCAATCTT ACCATTTCTA TITATTCTTG AGAAGTTACT TAATGCTGAA | 7080 |
| | CGTTTAACGA CACCACGTTT AGTTGCAAAC ACTAAGAAGT TGTCTTCACT TTCAAGGTCT | 7140 |
| 15 | TTAACAGCAA TCATTGTACT AATGACTTCA TCATTTTCAA GTTCAATAGC ATTCACTACA | 7200 |
| | GGAATACCTT TAGACTGTCT TGATAACTCA GGCACTTCGT AACCTTTAAG TTTGTATACA | 7260 |
| | CGACCTTTGT TAGTAAAGAA CAATACATGG TCATGTGTAC TTAAAGTTAC CAATTGACTG | 7320 |
| 20 | ACAAAATCTT CTTCCAATGT ATTCATACCT TGAACACCAC GACCACCACG GTTTTGAGCA | 7380 |
| | CGATATGTAG ATACCGGCAA ACGTTTAATG TAGTTATTAT GGCTTAGTGT AATTACTATT | 7440 |
| | TGTTCTTCTG GAATTAAGTC TTCGTCCTCT AAGTCTTCAA ATCCACCTAA TTGAATTTCT | 7500 |
| 25 | GTACGACGAT CATCACCGAA ACGATCTCTA ATTTCAGTCA ATTCATCTCT AACTAACTGT | 7560 |
| | AATAACACTT CTTCATCAGC TAAGATTGCT TCTAATTCAC TAATAAATT TAATAACTCA | 7620 |
| | TTATATTCAG CTTCAATTTT GTCTCTCTT AAACCTGTTA GACGTCTTAA ACGCATGTCT | 7680 |
| 30 | AAAATAGCTT GAGCTTGTTT TTCAGAAAGT TTGAAGCGTT GTTGCAAGCT TTCCATTGCA | 7740 |
| | ACTITATETG TATETGACTE ACGAATEGTT GAAATAATTT CATEGATATG GTCAAGTGCG | 7800 |
| | ATACGTAATC CTTCTAAAAT GTGGGCACGA TCTTTAGCTT TACGTAAGTT GTATTGCGTA | 7860 |
| 35 | CGTCTTCTAA CAACTGTCTT TTGATGCTCT AAATAATGTA CCAACGCTTC TTTTAAATTA | 7920 |
| | ATAAGCTTCG GTCTACCATT TACAAGTGCA ATCATATTCA CACCAAATGA TGTTTGAAGA | 7980 |
| 40 | GGTGTTTGTT TGTATAAGTT ATTTAAAATG ACACTAGCAT TTGCATCCTT ACGCACATCA | 8040 |
| <u> </u> | ATAACGACAC GCACACCAGT ACGTAAACTT GTTTCATCAC GTAAATCAGT GATACCGTCA | 8100 |
| | ATTTTCTTGT CACGAACGAG CTCTGCAATT TTTTCAATCA TACGAGCCTT ATTCACTTGG | 8160 |
| 4 5 | AAAGGAATTT CAGTGACAAC AATACGTTGA CGTCCGCCTC CACGTTCTTC AATAACTGCA | 8220 |
| | CGAGAACGCA TITGAATTGA ACCACGACCT GTTTCATATG CACGTCTAAT ACCACTCTTA | 8280 |
| | CCTAAAATAA GTCCAGCAGT TGGGAAATCA GGACCTTCAA TATCCTCCAT TAACTCAGCA | 8340 |
| 50 | ATTGAAATAT CAGGGTTCTT ACTTAAGCTA AGTACACCAT TGATTAATTC TGTTAAGTTA | 8400 |
| | TGTGGTGGAA TATTCGTTGC CATACCTACC GCGATACCTG ATGCACCATT CCCTAATRAC | 2462 |

| | AAATCTATTG | TATCTTTATT | AATATCACGT | AACAGTTCAA | GTGTGATTTT | AGTCATACGC | 8580 |
|----|-------------------------|----------------|----------------|--|------------|--------------|-------|
| | GCTTCAGTAT | AACGCATTGC | TGCTGCGCCA | TCTCCATCCA | TTGAACCAAA | GTTACCTTGG | 8640 |
| 5 | CCATCAACAA | GCGGATAACG | ATAACTGAAA | TCTTGAGCCA | TACGTACCAT | TGCTTCATAA | 8700 |
| | ATAGATGAGT | CACCATGAGG | GTGATATTTA | CCCATTACGT | CACCAACGAT | ACGTGCTGAT | 8760 |
| | TTTTTATATG | ATTTATCCGG | TGTCATACCT | TGTTCATTTA | ATCCATATAG | TATACGACGA | 8820 |
| 10 | TGTACTGGTT | TTAAACCGTC | ACGAACATCT | GGCAATGCAC | GAGCAACGAT | AACACTCATC | 8880 |
| | GCATAATCTA | AAAATGATTC | ACGCATTTCA | CTGGTAATAT | TTCGTTCATT | TATTCTTGAT | 8940 |
| 15 | TGAGGTAATT | CAGCCATCAA | GAGTTCCTCC | TTCAAAAGTT | CAGTTCACAG | CGCTTAGAAG | 9000 |
| 13 | TCTAAGTTTG | CATAAACTGC | ATTATCTTCT | ATAAATTGTC | TACGGTTTTC | TACAACGTCA | 9060 |
| | CCCATTAACA | TTTCAAATGT | TTGGTCCGCT | TCAATCGCAT | CTTCAAGTTT | TACTTGTAAA | 9120 |
| 20 | AGAGCGCGGT | GCTCAGGGTT | CATTGTTGTT | TCCCALAATT | GATCTGCATT | CATTTCTCCA | 9180 |
| | AGACCTTTGT | ATCGTGCAAT | AGACCATTTT | GGTGTTGGAT | TCAATTCAGA | TTTAAGTTTA | 9240 |
| | TCAAGTTCCC | TATCATTGTA | TACATAATAC | TTTTGTTTAC | CTTGTGTCAG | TTTATACAAC . | 9300 |
| 25 | GCTGGCTGTG | CAATATACAC | ATAGCCTGCT | TCAATTAACG | GTCTCATAAA | TCGATAGAAG | 9360 |
| | AATGTTAATA | ACAATGTTCT | AATATGCGCT | CCATCCACAT | CGGCATCAGT | CATAATGACG | 9420 |
| | ATTTTGTGAT | ATCTTGCTTT | CGCTAGATCA | AAGTCGCCAC | CGATTCCTGT | ACCAAATGCT . | 9480 |
| 30 | GTGATCATTT | GACGAATTTC | ATTGTTATTC | AAAATTCTAT | CTAATCGTGC | TTTTTCAACA | 9540 |
| | TTTAATATCT | TACCTCGTAA | TGGTAAAATC | GCCTGCGTTC | TAGAGTCACG | ACCAGATTTT | 9600 |
| | GTAGACCCCC | CGGCAGAGTC | CCCTTCGACT | AAGAAAATCT | CACATTCTTC | AGGACTTTTA | 9660 |
| 35 | CTAGAGCAAT | CGGCTAATTT | ACCTGGAAGG | CTTGCTACAT | CTAACGCTGA | TTTACGACGT | 9720 |
| | GTTACTTCAC | GCGCTTTTTT | CGCAGCAACA | CGTGCACGTG | CCGCCATAAT | ACCTTTTTCA | 9780 |
| | ACCACTGTAC | GTGCGACTTG | TGGATTTTCA | TATAAAAATC | GTTCAAAGTG | CTCTGAGAAT | 9840 |
| 40 | AATTTATCTA | CAACTTGACG | CACTTCAGAA | TTACCTAATT | TTGTCTTCGT | TTGACCTTCG | 9900 |
| | AATTGAGGAT | CACCATGTTT | GATAGATATA | ATTGCTGTCA | TACCTTCACG | TGTATCTTCA | 9960 |
| | CCAGAAAGTC | TATCTTTTTC | TTCTTTCATA | ATCTTGCTAC | TTAAACCATA | ACTATTTAAG | 10020 |
| 45 | ACACGCGTTA | ATGCACGTTT | GAATCCGTCT | TCATGCGTAC | CACCTTCATA | CGTATGAATG | 10080 |
| | TTATTTGCGT | aagttaaaag | ATTTGTGGCA | TATCCTGAGT | TATATTGAAT | CGCAATTTCT | 10140 |
| 50 | ACTTCAATAT | CATCTTTAGA | TTGATGAATA | TAAATTGGCT | CATCATGAAT | AGGTTCTTTA | 10200 |
| | THE THE PERSON NAMED IN | ATA A CTCA A C | COTA CCA MOVEA | ************************************** | CATACTCATA | CCACTCTTCT | 1026 |

| | GCAAGCTCTC | TAATACGCTG | CTGTAATGTT | TCATAGTTGT | ATACAGTTGI | CTCTGTGAAG | 10380 |
|----|------------|------------|------------|--|------------|------------|-------|
| - | ATTTCTCCAT | CTGCTTTAAA | ACGAAtGaCA | GTACCTGTCT | TATCAGTEGT | GCCAACTTCT | 10440 |
| 5 | TTTAAGTCAA | ATTGAGGTAC | ACCTITTTA | TATGCTTGAT | GATATATAGT | CTCATTTCTG | 10500 |
| | TGTACATATA | CTTCTAAGTC | TTGTGACAAT | GCGTTTACAA | CTGATGAACC | AACACCATGT | 10560 |
| 10 | AAACCACCAG | ATACTTTGTA | TCCGCCACCG | CCAAATTTAC | CACCAGCATG | TAAAACAGTT | 10620 |
| 70 | AAAATAACTT | CGACAGCTGG | ACGTCCCATT | TTTTCTTGAA | TATCAACTGG | GATACCACGT | 10680 |
| | CCGTTATCCG | TTACTTTAAT | CCAGTTATCT | TTTTCAATAA | CAACTTCAAT | TTGATTTGCA | 10740 |
| 15 | TAACCAGCTA | ATGCTTCATC | GATACTATTA | TCGACAATTT | CCCACACTAA | ATGGTGCAAA | 10800 |
| | CCTCTCTCTG | AAGTCGATCC | TATATACATA | CCTGGTCTTT | TACGTACTGC | TTCTAAACCT | 10860 |
| | | | | | TGTTTACATC | | 10920 |
| 20 | | | | | TAATACGATA | | 10980 |
| | | - | | | TAGTGACAAA | | 11040 |
| | | | • | • | CATCTAATTC | | 11100 |
| 25 | | | • | | TCATTAACTC | | 11160 |
| | AATTTAATGG | ACAAAGCCGT | TGTACGTTGC | TGTCCTTGAG | AACCATATGT | TTGAGCATCC | 11220 |
| _ | | | ŕ | | CGAATAAGCT | | 11280 |
| 30 | | | | | TTTCTTCAAG | | 11340 |
| | | | | | TCAGCGACAA | | 11400 |
| | | | | | CTAGCTCTTG | | 11460 |
| 35 | | | | | GCTGATTTAA | | 11520 |
| | | | | | AGTAATTATT | | 11580 |
| | • | | | | CAGAAATTTG | | 11640 |
| 40 | | | | | CAATATTCAA | · | 11700 |
| • | | | | | GACGACTTTG | | |
| | | *. | | | TGTTAATGGG | , | 11760 |
| 45 | TGT | | | ************************************** | IGIIMAIGGG | CAICGIGCEG | 11820 |
| | | | • | | | | 11823 |

(2) INFORMATION FOR SEQ ID NO: 137:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 692 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

55

| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 137: | |
|----|--|-----|
| _ | ATAATTATTA ACATGGTGTG TTTAGAAGTT ATCCACGGCT GTTATTTTTG TGTATAACTT | 60 |
| 5 | AAAAATTTAA GAAAGATGGA GTAAATTTAT GTCGGAAAAA GAAATTTGGG AAAAAGTGCT | 120 |
| | TGAAATTGCT CAAGAAAAAT TATCAGCTGT AAGTTACTCA ACTTTCCTAA AAGATACTGA | 180 |
| 10 | GCTTTACACG ATTAAAGATG GTGAAGCTAT CGTATTATCG AGTATTCCTT TTAATGCAAA | 240 |
| | TTGGTTAAAT CAACAATATG CTGAAATTAT CCAAGCAATC TTATTTGATG TTGTAGGCTA | 300 |
| | TGAAGTTAAA CCTCACTTTA TTACTACTGA AGAATTAGCA AATTATAGTA ATAATGAAAC | 360 |
| 15 | TGCTACTCCA AAAGAAACAA CAAAACCTTC TACTGAAACA ACTGAGGATA ATCATGTGCT | 420 |
| | TGGTAGAGAG CAATTCAATG CCCATAACAC ATTTGACACT TTTGTAATCG GACCCGGTAA | 480 |
| | CCGCTTTCCA CATGCAGCGA GTTTAGCTGT GGCCGAAGCA CCAGCCAAAG CGTACAATCC | 540 |
| ?0 | mTTATTTATC TATGGAGGTG TTGGtTTAGG aAAAACCCAT TTAATGCATG CCATTGGTCA | 600 |
| | TCATGTTTTA GATAATAATC CAGATGCCAA AGTGATTTAC ACATCAAGTG AAAAATTCAC | 660 |
| | AAATGAATTT ATTAAATCAA TTCGTGATAA nA | 692 |
| ?5 | (2) INFORMATION FOR SEQ ID NO: 138: | |
| 30 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 7900 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 138: | |
| 35 | ATACTGTAGC GCAAATTTCA CAATGGCATG TTATAGAAGA TTTAGTTACG AATGAATTAG | 60 |
| | GTATTAGTAT TTTACCAACA TCAATTTCAG AGCAACTAAA TGGAGATGTG AAGCTGLACG | 120 |
| | CATTGAAGAT GCTCATGTAC ATTGGGAATT AGGTGTTGTT TGGAAGAAGG ATAAACAATT | 180 |
| 10 | AAGTCATGCC ACAACGAAAT GGATAGAATT TTTGAAAGAC CGTTTAGGCT AACATATTAA | 240 |
| | TAAAGCACTC ATTATTTAAG GCGCATCATT ACGTGGGTCA TTGAAATAAT GAGTGTTTTT | 300 |
| 15 | TTGTGAAAAT GAAGTGAAAT TTAGAGAGCG TTTCCATAGA AAATAGTAAT ACAAACTATA | 360 |
| | AAAAAAGAGT ATTTTTATAT TGTGTACGCC ATCTTTATAA TAGTTATTGT AACAATTTAG | 420 |
| | ACATATTIAG ARAGGGATGG CGCCATGCAC ARAGTCCART TARTARTCAR ACTACTACTA | 480 |
| 50 | CAACTAGGAA TCATCATTGT GATTACTTAT ATTGGCACAG AAATTCAAAA GATTTTTCAT | 540 |

55

| | ATTGTACCGC | TAACTTGGGT | AGAAGACGGT | GCAAACTTTT | TATTAAAGAC | GATGGTCTTT | 660 |
|----|------------|------------|--------------|-------------|------------|----------------|------|
| | TTCTTCATAC | CGTCAGTTGT | ' AGGLATTATG | GaTGtgCTTC | CGAAATTACG | CTAAATTATA | 720 |
| 5 | TACTCTTTTT | CGCAGTCATT | ATCATAGGAA | CATGTATCGT | TGCATTATCT | TCAGGTTATA | 780 |
| | TTGCTGAAAA | AATGTCyGtT | AAACWTAAAC | ATCGTAAAGG | TGTAGACGCt | TATGAATGAT | 840 |
| | TACGTGCAAG | CCTTATTAAT | GATTTTGTTG | ACTGTCGTTT | TATATTATTT | CGCTAAAAGG | 900 |
| 10 | TTACAACAAA | AATATCCGAA | CCCATTTTTG | AATCCAGCAT | TAATTGCATC | TTTAGGAATT | 960 |
| | ATTTTTGTCT | TACTTATCTT | TGGAATTAGT | TATAACGGGT | ATATGAAAGG | TGGCAGTTGG | 1020 |
| 15 | ATCAACCATA | TTTTAAACGC | AACGGTCGTA | TGTTTAGCGT | ACCCACTTTA | TAAAAATAGA | 1080 |
| 15 | GAGAAAATTA | AAGACAATGT | CTCTATCATT | TTTGCAAGTG | TATTAACTGG | CGTCATGCTG | 1140 |
| | AATTTCATGT | TAGTGTTCTT | AACACTTAAA | GCATTTGGCT | ATTCTAAAGA | CGTCATTGTA | 1200 |
| 20 | ACGTTATTGC | CCCGATCTAT | AACAGCCGCA | GTAGGTATCG | AAGTGTCACA | TGAACTAGGT | 1260 |
| | GGTACAGATA | CGATGACCGT | ACTITITATT | ATCACAACGG | GTTTAATCGG | TAGTATTTTA | 1320 |
| | GGTTCGATGT | TATTAAGATT | TGGAAGATTT | GAATCTTCTA | TCGCCAAAGG | ATTAACGTAT | 1380 |
| 25 | GGGAATGCGT | CACATGCATT | TGGCACAGCT | AAAGCACTAG | AAATGGATAT | TGAATCCGGT | 1440 |
| | GCATTTAGTT | CAATTGGGAT | GATTTTAACT | GCAGTTATTA | GTTCAGTGTT | AATACCTGTT | 1500 |
| | CTAATTTTAT | TATTCTATTA | ATTTAGATAT | TTAAAATGAT | AGACAGAAAG | GGAGGCTATT | 1560 |
| 30 | AGTAATAATG | GCAAAAATAA | AAGCAAATGA | AGCATTAGTT | AAAGCATTAC | AAGCaTGGGA | 1620 |
| | TATAGATCAC | TTGTATGGTA | TTCCAGGAGA | CTCAATCGAC | GCATAGTCGA | TAGTTTACGT | 1680 |
| | ACAGTGAGAG | ATCAATTTAA | ATTTTATCAT | GTACGTCATG | AAGAAGTAGC | AAGCTTAGCG | 1740 |
| 35 | GCTGCTGGTT | ACACAAAATT | AACTGGTAAA | ATCGGTGTGG | CATTAAGTAT | CGGTGGCCCT | 1800 |
| | GGTTTAATTC | ATTTATTAAA | TGGTATGTAT | GATGCCAAAA | TGGATAATGT | ACCGCAATTA | 1860 |
| | ATATTATCTG | GACAAACGAA | TAGTACAGCA | CTTGGAACGA | AAGCATTCCA | AGAAACAAAT | 1920 |
| 40 | TTACAAAAAT | TATGTGAAGA | TGTAGCCGTT | TATAATCACC | AAATTGAAAA | AGGTGACAAT | 1980 |
| | GTGTTTGAAA | TCGTTAACGA | AGCAATTCGT | ACGGCATATG | AACAAAAAGG | TGTAGCTGTT | 2040 |
| | GTTATTTGTC | CTAACGACTT | ATTAACTGAA | AAAATTAAAG | ATACAACGAA | TAAACCAGTA | 2100 |
| 45 | GATACATCAA | GACCAACAGT | AGTATCACCA | AAATATAAAG | ACATCAAAAA | AGCGGTTAAA | 2160 |
| | CTAATTAATA | AAAGTAAAAA | GCCTGTCATG | TTAATTGGTG | TAGGTGCGAA | ACATGCGAAA | 2220 |
| 50 | GATGAGCTAC | GTGAATTTAT | TGAAATGGCT | AAAATTCCTG | TCATTCATTC | ATTACCAGCT | 2280 |
| 50 | AAAACAATCT | TGCCGGATGA | TCATCCATAT | AGTATCGG+ A | астрасства | 3 3 TCCCT 3 CC | 2240 |

| | CCATATGTGG | ATTACTTACC | TAAGAAAAAT | ATTAAAGCCA | TTCAAATTGA | CACAAATCCT | 2460 |
|-----------|------------|------------|------------|------------|------------|---|------|
| | AAAAATATCG | GACATCGTTT | CAATATTAAT | GTAGGAATTG | TTGGAGATAG | TAAAATTGCG | 2520 |
| 5 | TTGCATCAGT | TAACTGAAAA | TATTAAACAT | GTTGCTGAAA | GACCATTCTT | AAACAAAACG | 2580 |
| | TTAGAACGTA | AAGCGGTTTG | GGATAAATGG | ATGGAACAAG | ATAAAAATAA | TAATAGTAAA | 2640 |
| | CCATTACGTC | CAGAACGATT | AATGGCATCA | ATCAATAAAT | TTATTAAAGA | TGATGCAGTG | 2700 |
| 10 | ATTTCAGCAG | ATGTAGGTAC | AGCAACAGTT | TGGTCAACTC | GATACTTAAA | CCTTGGTGTA | 2760 |
| | AATAACAAGT | TCATCATTTC | AAGTTGGTTA | GGTACAATGG | GTTGCGGTCT | TCCAGGTGCA | 2820 |
| | ATTGCATCAA | AAATTGCATA | TCCAAATAGA | CAAGCCATCG | CAATTGCTGG | TGACGGTGCA | 2880 |
| 15 | TTCCAAATGG | TAATGCAAGA | CTTCGCTACA | GCAGTACAAT | ATGATTTACC | TTTAACTGTA | 2940 |
| | TTTGTACTTA | ATAACAAACA | GTTAGCATTT | ATTAAATATG | AACAACAAGC | AGCTGGTGAA | 3000 |
| | TTAGAATATG | CAGTTGATTT | TTCTGATATG | GATCATGCAA | AATTTGCTGA | GGCAGCAGGT | 3060 |
| 20 | GGTAAAGGTT | ATACAATTAA | GAGTGCTAGC | GAAGTÄĞATG | CTATAGTCGA | AGÁGCATTA | 3120 |
| | GCACAAGATG | TACCAACGAT | TGTAGATGTA | TATGTTGATC | CTAÀTGCTGC | GCCATTACCA | 3180 |
| 25 | GGTAAAATTG | TAAATGAAGA | AGCGCTTGGT | TATGGTAAGT | GGGCATTTAG | ATCAATTACT | 3240 |
| | GAAGATAAAC | ATTTAGATTT | AGATCAAATT | CCACCAATTT | CAGTGGCAGC | AAAACGTTTC | 3300 |
| • | TTATAACTGA | TTTAAAGGTT | ATCACAATTG | AATTGAACTA | TAAAAACGGT | AATTTCTATT | 3360 |
| <i>30</i> | TCAACAAAAT | GGGAATTGCC | GTTTTGTTTA | TTTATCACAA | ATGATCGTAC | TGAATTGATG | 3420 |
| | ATAAAATTGT | GAAAAAGTTG | TTGAAAACGC | TTTTACAAAT | ATGTATAATA | GCTATGAATT | 3480 |
| | AGATATCACT | TGCGTGTTAC | TGGTAATGCA | GGCATGAGCA | AACAACCGCA | CTATGAGAAT | 3540 |
| 35 | AGTCTTGTTT | GTTCATGCCT | GCTTTTTTTG | TACATGGAAG | CGGAAATTGA | GATAGGGGAT | 3600 |
| | GTTTATATGT | TTAAGAAATT | GTTTGGACAA | TTGCAACGTA | TCGGTAAAGC | ATTAATGTTA | 3660 |
| | CCTGTTGCGA | TTTTACCAGC | AGCTGGTATT | TTATTAGCGT | TTGGTAACGC | AATGCACAAC | 3720 |
| 40 | GAACAATTAG | TAGAAATTGC | ACCATGGTTA | AAAAACGATA | TCATTGTAAT | GATTTCGTCG | 3780 |
| | GTCATGGAAG | CAGCAGGACA | AGTTGTATTT | GATAACTTGC | CATTATTATT | TGCAGTTGGT | 3840 |
| • | ACAGCACTTG | GATTAGCAGG | AGGAGACGGT | GTTGCAGCAT | TAGCAGCGCT | AGTAGGTTAC | 3900 |
| 45 | TTAATTATGA | ATGCAACAAT | GGGGAAAGTG | TTGCACATTA | CAATTGATGA | CATTTTCTCA | 3960 |
| | TATGCCAAAG | GGGCAAAAGA | ATTAAGTCAA | GCAGCGAAAG | AACCAGCACA | TGCTTTAGTA | 4020 |
| | TTAGGTATTC | CAACGTTACA | AACGGGTGTG | TTTGGTGGTA | TTATCATGGG | TGCTTTAGCC | 4080 |
| 50 | GCATGGTGTT | ACAACAAATT | TTATAATATT | ACACTACCAC | CATTITUTE | ארידייייייייייייייייייייייייייייייייייי | 4140 |

| | AGCTTTGCG | r ggccaccaat | TCAAGATGG | A TTAAATAGTT | TATCGAATT | CTTATTAAAT | 426 |
|----|------------|--------------|------------|--------------|------------|------------|--------|
| • | AAAAATTTAA | A CATTAACAAC | GTTTATATT | C GGTATTATTG | AACGCTCATI | AATTCCATTT | 432 |
| 5 | GGTTTACATO | ATATTTTCTA | TTCACCGTT | TGGTTTGAAT | TCGGAAGTTA | TACAAATCAC | 4386 |
| | GCAGGTGAAT | TGGTTCGTGG | TGACCAACGT | T ATTTGGATGG | CACAATTGAA | AGATGGCGTA | 444(|
| • | CCATTTACTO | CTGGTGCATT | TACTACTGGT | TAAATATCCAT | TTATGATGTT | TGGTTTACCA | 4500 |
| 10 | GCGGCGCAT | TTGCTATTTA | TAAAAATGCA | CGACCAGAAC | GTAAAAAAGT | CGTGGGTGGT | 4560 |
| | TTAATGTTAT | CAGCAGGATT | AACTGCATTT | TTAACTGGTA | TCACTGAGCC | ATTAGAATTT | 4620 |
| | TCATTCTTAT | TTGTAGCACC | AGTACTTTAT | GGAATTCACG | TATTATTAGC | TGGTACATCA | 4680 |
| 15 | TTCTTAGTAA | TGCATTTATT | AGGCGTTAAA | ATTGGTATGA | CATTCTCAGG | TGGTTTCATA | 4740 |
| | GATTATATTT | TATATGGTTT | ATTAAACTGG | GATCGTTCAC | ACGCATTATT | AGTTATTCCA | 4800 |
| 20 | GTCGGTATTG | TATATGCTAT | CGTGTATTAC | TTCTTATTCG | ACTTTGCAAT | TCGTAAGTTT | 4860 |
| 20 | AAATTGAAAA | CACCAGGTCG | TGAAGATGAA | GAAACTGAAA | TTCGTAACTC | TAGTGTCGCA | 4920 |
| | AAATTACCAT | TTGATGTCTT | AGATGCAATG | GGTGGAAAAG | AAAACATTAA | ACATTTAGAT | 4980 |
| 25 | GCATGTATTA | CACGTCTACG | CGTAGAAGTG | GTTGATAAAT | CAAAAGTAGA | TGTAGCAGGT | 5040 |
| | ATTAAAGCTT | TAGGCGCATC | AGGTGTATTA | GAAGTTGGAA | ACAATATGCA | AGCTATCTTT | 5100 |
| | GGTCCAAAAT | CAGATCAAAT | TAAACATGAT | ATGGCCAAGA | TTATGAGTGG | TGAAATTACG | . 5160 |
| 30 | AAACCAAGTG | AAACGACAGT | GACTGAAGAA | ATGTCAGATG | AACCAGTTCA | CGTAGAAGCA | . 5220 |
| | CTTGGAACAA | CAGACATCTA | TGCACCAGGT | ATCGGTCAAA | TCATTCCATT | ATCAGAAGTA | 5280 |
| | CCTGATCAAG | TATTCGCTGG | TAAAATGATG | GGTGATGGTG | TTGGCTTTAT | CCCTGAAAAA | 5340 |
| 35 | GGTGAAATTG | TAGCACCGTT | TGATGGTACA | GTGAAAACAA | TCTTCCCTAC | GAAACATGCG | 5400 |
| | ATAGGATTAG | AATCTGAAAG | TGGCGTCGAA | GTACTTATTC | ATATTGGTAT | CGATACAGTG | 5460 |
| | AAACTGAATG | GTGAAGGATT | CGAAAGTCTG | ATTAACGTTG | ATGAAAAAGT | AACACAAGGT | 5520 |
| 40 | CAACCATTAA | TGAAAGTGAA | TTTAGCATAC | TTGAAAGCAC | ACGCACCAAG | CATCGTTACA | 5580 |
| | CCAATGATTA | TTACAAATCT | TGAAAATAAA | GAACTTGTCA | TTGAAGATGT | ACAAGATGCT | 5640 |
| | GATCCAGGTA | AGCTAATTAT | GACAGTCAAA | TAATGATTAA | AAATGAAACA | GCATATCAAA | 5700 |
| 45 | TGAATGAACT | TTTAGTCATT | CGTAGTGCGT | ATGCGAAGTA | GCGAGTTGAA | AGAGAATACG | 5760 |
| | TTACAAAAGG | CAGTAGCTTA | AAATGAAGCT | ACTGCCTTTT | TAGTGCGCAA | TGATGTATAG | 5820 |
| | CAGGTGTGTT | GATGITAATA . | AGTTAAATAT | TAGTGTTAGA | TATAGAAAAC | ATTGCTTATG | 5880 |
| 50 | TTTTTGTCAC | ATTITAGAAA | AATGCATCTT | CGCGACTAGC | САААТТААТА | GTCTCATTCA | .E940 |

| | AATAAATTAA | CATGATTTTA | AATCTATTTG | TAAGATAAGG | AGATTTGTCA | TTATGACAAC | 6060 |
|------------|------------|------------|------------|------------|-------------------|------------|------|
| | AGAAGGTCTA | TTAGTTGCAG | AGAAAGAAAT | CGAAGTGAAT | GGTTACGACA | TTGATGCGAT | 6120 |
| 5 | GGGTGTCGTT | AGTAATATCG | TTTATATTAG | ATGGTTCGAA | GATTTGAGAA | CAGCGTTTAT | 6180 |
| | TAATCAGCAC | ATGAATTACT | CAACAATGAT | CAATCAAGGC | ATTTCACCTA | TACTTATGAA | 6240 |
| | AACGGAAGCA | GAGTATAAAG | TACCTGTCAC | AATACATGAC | AAACCAGTAG | GTCGTATTTA | 6300 |
| 0 | CTTAGTTAAA | GCAAGCAAGA | TGAAATGGGT | GTTTCAGTTT | GAAATTGTGT | CCGCACATGG | 6360 |
| | CGTGCATTGT | ATTGGTACAC | AGACAGGCGG | TTTTTACAGA | TTGAGTGATA | AGAAGATAAC | 6420 |
| | CTCTGTGCCA | CAAGTGTTTC | AAGACATTTT | AGCAACAAAA | TAATGACTTC | TAAAATTTTA | 6480 |
| 5 | ATAAAAAGTA | AGAAGGTGTT | CGAAATGGTT | AAGCAATTAA | ATAGTGTCGA | AGCATTCCGT | 6540 |
| | GAATTTATTC | ATCAATATCC | GTTAGCAGTT | GTACATGTCA | TGCGCGATCA | GTGTAGCGTG | 6600 |
| | TGTCATGCCG | TTTTACCACA | AATTGAAGAC | TTGATGCAAT | CATATCCCAA | TGTGCCATTA | 6660 |
| 20 | GCTGTGATTA | ATCAAAGTCA | GGTGGAAGCT | ATTGCTGGAG | AATTAAATAT | TTTCaCTGTA | 6720 |
| | CCTGTGGATT | TATTTTAT | GAATGGAAAA | GAAATGCATC | GTCAAGGGCG | TTTTATCGAT | 6780 |
| ?5 | ATGCAACGTT | TTGAACATCA | TCTTAAGCAA | ATGAATGATA | GTGTAAATAA | CGATGTCGAT | 6840 |
| .0 | GAGCATTAAT | ATCGCAAATG | ATTAGCATTG | CTAAGATTAT | GTAGACATCA | TAACTTATTT | 6900 |
| | CCCAGTAAAT | ATTGGTAGTA | ATTAGAATCA | GCATGGTACA | GTAGAACTAT | AGTAGAAATC | 6960 |
| 30 | ATCAAAGAGG | AGTGACGACA | AATGCGTAAA | AAATGGTCTA | CACTTGCGTT | TGGATTTTTA | 7020 |
| | GTTGCAGCAT | ACGCACATAT | TAGAATTAAA | GAAAAACGCA | GTGTGAAAAG | TTATATGTTA | 7080 |
| | GAACAAGGTA | TACGATTATC | TAGAGCTAAG | CGTCGTTTTA | TGTATAAAGA | AGAAGCGATG | 7140 |
| 35 | AAAGCATTAG | AAAAAATGGC | GCCACAGACA | GCAGGCGAAT | ATGAGGGAAC | CAATTATCAG | 7200 |
| | TTTAAGATGC | CAGTAAAAGT | GGATAAGCAC | TTCGGTTCAA | CCGTTTATAC | CGTTAACGAT | 7260 |
| | AAAĊAAGATA | AGCATCAACG | CGTTGTATTA | TATGCACATG | GAGGCGCATG | GTTCCAAGAC | 7320 |
| 40 | CCACTCAAAA | TTCATTTCGA | ATTTATTGAT | GAACTTGCAG | AAACACTCAA | TGCTAAAGTC | 7380 |
| | ATCATGCCAG | TATATCCGAA | GATTCCGCAT | CAAGATTATC | AAGCGACGTA | TGTGCTTTTT | 7440 |
| | GAAAAGTTGT | ACCATGATTT | ATTGAATCAA | GTAGCAGATT | CTAAACAAAT | CGTTGTAATG | 7500 |
| 4 5 | GGTGACTCTG | CGGGCGGTCA | AATTGCTTTA | TCATTTGCTC | AATTGTTAAA | AGAAAAACAT | 7560 |
| | ATTGTGCAAC | CAGGACATAT | TGTATTAATT | TCACCAGTTT | TAGATGCAAC | GATGCAGCAT | 7620 |
| | CCTGAAATTC | CTGACTACTT | AAAGAAAGAC | CCAATGGTAG | GTGTGGATGG | CaGTGTGTTC | 768 |
| 50 | mm.comcc | ********* | 001010100 | ********** | 3 C 3 3 3 CT 3 TC | ACCARTTART | 774 |

| | CCAGATGCTT TGAACTTATC GCAATTGTTG AGTGCGAAAG GTATCGAACA TGACTTTATA | 7860 |
|----|--|--------|
| | CCTGGATATT ACCAATTCCA TATTTATCCA GTATTTCCGA | 7900 |
| 5 | (2) INFORMATION FOR SEQ ID NO: 139: | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1984 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 139: | |
| 15 | GTCTAAATAA ACAAAATTAT CATTGATTAC TGAACTGGCA TTTCGAAGTA ATGCTTCAAT | 60 |
| | ATCATTCGAA TATTTCTTCA ATTTATGATT GTGAAATAAT TCTTGCATCA AAAATGGTCT | 120 |
| 20 | TTGGTCACAT GAATGTGCAT CTGAAGCTAC AAAATGAGCC AAATTACATT CTATAAATTG | 180 |
| 20 | TAATGATAAC TTTTGAATGT TTTTACCAAA TCCACCAACT AAAGAACTCG ATGTTAATTG | 240 |
| | ACTCAGTGCC CCATTTGCAA CCAATTCATA TAATATTTCC GGATTTTTGG CGATACTTCT | 300 |
| 25 | ATTTCTTTCA GGATGTGCAA TGATTGGTAT GTAACCTCTC GATTGTATTT CAAAAAACAA | 360 |
| | TIGTTTTGTA TAATGTGGTA CTTCGCCCGT TGGAAATTCA ATTAATAAAT ATTTCGAACG | 420 |
| | ATTAATACCT TGAATACTAC CATTATCTAA GCCTTTCAGA ATCGAATCTG TAATTCTAAT | 4,80 |
| 30 | TTCTTGCCCG GGAAATAATT TAATATCCAA TGCTTGAACT TCTGGATGCG TTCTTAACTC | 540 |
| | CGCCAATTTC ACAAGCACTT GTTGAAATGT ATTATCATAT CTCGGATGCA AATGATGAGG | 600 |
| | TGTCGCTACA ATACTTGTTA CACCTTCATC CTTAGCTTGC TTTAATAGTG CAATACTCTT | 660 |
| 35 | TTCAATTGTT TTAGGACCAT CATCTATATC AACTAATATA TGGTTATGAA TATCAATCAT | 720 |
| | GATTCATCAG TCCCATAATA TGCATAGTAA CTAGCACTTT TATCTTTAGG CATTCTATTT | 780 |
| | AAGACTACAC CTAATAATTT AGCACCTGTT GCTTCAATAA GTTCTTTTCC TTTTTTAACT | 840 |
| 40 | TCATCTCTAT TATTATTTTC CGAATTAACT ACGTAGACAA CATTGCCGGT AAACTTTGAA | 9.0.0. |
| | AATAATTGCG CATCTGTAAC TGTGTTCACT GGTGGCGTAT CGATAATTAC AAAGTTATAA | 960 |
| | TTCATCAATA ATGTGTCATA CAAATTTGCA AATGCCCTTG ATGTAATTAA CTCTGACGGA | 1020 |
| 45 | TTCGGTGGGA TTGGCCCAGA CGTCAAGACG TCTAAATCTT GAATTTCAGT TGAGATAATA | 1080 |
| | CTGTCTTGAT AAGTTGACCA ATTTAGCAAT AAACTTGATA GGCCTTCATT GTTTGGCAAA | 1140 |
| 50 | TTAAAAATAT AATGCTGCGT AGGTTTACGC ATATCCCCGT CTACGATTAG TGTTTTATAA | 1200 |
| | CCTGCTTGCG CATATGCAAC TGCTAAATTT GCTGCAATTG TAGACTTACC TGCGCCTGGT | 1260 |

55

| GATCTTATGC | CTCGAAATTT | CTCGCTAATA | GGTGACTTTG | GTTGTTCATG | GACAATTAAA | 1380 |
|------------|------------|------------|------------|------------|------------|------|
| CTTGATGTAC | TTCYTCGTGT | ATTCGTCATG | GTAATTCCTC | GTAAATTAAA | ATTTTTGTAT | 1440 |
| TGAACCTAAA | ATAGGTAATC | CTAGTTGCGA | TTCAACATCT | TCTTCTGTCT | TAATACGCTT | 1500 |
| ATCTAATAAT | TCTTTTAAGA | AAATAATCAA | TATTGCTAAA | ACAATACCAA | CAATAATGCT | 1560 |
| GATAACTAAG | TTGACAGATA | CTATTGGAGA | TACTTTTACA | GCATTATCAT | GTGCTGAGGA | 1620 |
| AAGTATCGTA | ACATTATCAA | CACTCATAAT | TTTAGGCATG | TCATGAGCAA | AAACTTTAGA | 1680 |
| TATTTTATTA | ACAATTTTGT | CAGATTCAGA | TTTATTCCCA | GTGGTAACTG | ATACAGTAAT | 1740 |
| AATTTGAGAG | TTTGTTTGAT | TGGTTACTTT | TAAAAATGAA | TTCAACTCAG | CTGTTGAATA | 1800 |
| CTGACCATCA | Anttetetag | ATACTTTATC | TAGAATTCTA | GGACTTTTGA | TAATTTCCGT | 1860 |
| ATATGTATTA | ACAGACTGCA | AACTACTTTG | AACATTTTGG | AAAGCTAAAT | CACTTGAGGA | 1920 |
| CTTTTTCATG | TTCACTAATA | TTTGAGTAGA | AGCAGTATAT | TTGTCAGGCA | TAACAAAAA | 1980 |
| GGTT | | | | | | 1984 |

(2) INFORMATION FOR SEQ ID NO: 140:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6272 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 140:

| CAAATCCCTT | GGTGATGALA | AAtGtATTGC | TGTGTAGCCA | AATAATCTTC | GTATATATGA | 60 |
|------------|-------------|------------|------------|------------|------------------|-----|
| CTGACGTTCA | ACAACAGCTT | GCAATCGTTT | CGTTGGTACA | GTTACTTTCT | TCTTGTTAAA | 120 |
| GAÇACCATAT | TCAATTTTAA | GTTGCTCATT | TTCAAGCATC | ACCGAAAAGC | CATAAAATCT | 180 |
| TATCATTGTT | ATAATCGTTC | CAATAATATA | TGCCACTATT | AATACTAGTA | AAATGATGAT | 240 |
| TAATACTGAA | ATACTTACAA | TTTGAACCCA | TTGACTAATT | TCATGATTTA | GCTTCGACCA | 300 |
| TGGGATCAAC | TCTCTTACAG | CCCCGTAAAT | CGGTACTAAA | GCTGCTAACG | TTACACCAAT | 360 |
| GGCGCCACTG | GTCATTGCCA | TAAATAGTGA | TTCTTTAAAA | TTCATCTGAT | ATATAGGAAT | 420 |
| GCGTTTATTT | TTCTGATTAA | GCATACTATC | AGTGTTCTGC | ACTTCATCTA | AGCGACCTTC | 480 |
| TGCGATGTCT | TCCACATTAC | CTTCAATGTC | ATGATTACAG | TTGTCATTCT | TCTCAGCACT | 540 |
| AGACTTTTGC | GCCACTTCTG | TCTTCAACTC | TGTTTGCAAT | TGATCAATAT | ATCGTTCAAG | 600 |
| ATATTCACCT | TCTTTTTTTCC | ABATABCACT | TAACACAATA | CCNTCNCTTC | C-Trestatory & m | |

| | AATACGTTTT ATATTTAATT CTTTACGCTT TTTATTAAAA ATACCTGTTG TTAAAATGAA | 780 |
|-----------|---|------|
| 5 | ATAATTATCC tCAATCCAAT ATCGCGTGTT CATAATTCCG ACAATTTGAG AAATGTATGA | 840 |
| 5 | TATTAAAAAG AATACAAATA CAATACCTAT CCATAAATAT GATTCGGGAT TCGTATAATC | 900 |
| | AAAATCTTTC AATTGAAAGA TAATGAAAAT AAAAAAGACG ACTATGTTTT GTTTGATAGC | 960 |
| 10 | ATTGATTATG CCATTAAAAT ATGAAATCGG ATGTAATTTT TGAGGTTCAG ACATCACTTT | 1020 |
| | CAACCCCTCT CAAATTCGAC ATAGTTCTCT CTTCGATTAT TTTAACATCG TCATGAGACA | 1080 |
| | TCATCGGTAA ATAAATAGTA TGACCTGCAG TCATAAATCC AACTTTATAC AAATTAAGCA | 1140 |
| 15 | CTTTACTAAT TGGATTAGAT TTAATCGACA AGTATTGTAA ACGTTCAATT CGACTCGTTT | 1200 |
| | CTTCTTTATA TATAAAAAT GATGTACGAT ATTGTACACT TAGTTGATCA ACTTTATAAA | 1260 |
| | AGCGACAATG ATATTGCCAT AAAGGCTTAA TAAATAATTT TAATGTACTC AGAGCACCTA | 1320 |
| 20 | AAACCAACAA AATATAAAGT AAGTAATGTG GCCATTCAAA TCTTAACCAT ATAAAATAAA | 1380 |
| | AAATGACATA CACAGCTACA CTCAATATAA ATTCTAAGCC ATTCGTAATG TAGTAATACA | 1440 |
| 05 | ACAATGCTGA CTTAGGACTC TTAGTCAACT TAGTATAATC TGACATATAC CCCTCTCCCC | 1500 |
| 25 | AAATAAAAA TTATACGGAT TTATAATCTA TTTCATTTTA TTTTTATATG ATGATAATTA | 1560 |
| | TAGCATATGG AATATTTCAT GCTAATTTAT TCTTCCTAAA GGTACATCTA AAAATTTAAT | 1620 |
| 30 | TAAGCAGAAA GTGCTTGAAT TGCTAAAAAG ACACCATGTT ATAATTTTAT CAACATGATG | 1680 |
| | CCTTTCATCT ATAATCAATC TITCATCTTA TCAAGAGCGA TATTTAGTTC AAGCACATTC | 1740 |
| | ACATAATCAT TTGTTAACAC ACCACGCTGC TTACGATGTT GAATCAAGTC GGCCACTCTT | 1800 |
| 35 | GAAGTAGATA CATGACGAGC ATCAGCAATA CGAGGTGCTT GCTTCAATGC ATTTTCGACC | 1860 |
| | GTAATATGCG GATCTAAGCC CGACCCAGAA CTTGTTGCAG CATCTATTGT TACATTTGAA | 1920 |
| | TTCCCAAATT TAACATGATG TTTCATGCGT GCTATTAATT CGGTGTTTCC ATTCGATTCA | 1980 |
| 40 | TTACTTCCAC CTGAAGATAC GCCGTTTTTA TATAATTTTT CAGGATTCAT ATTATAATCA | |
| | ACTGCACTCG GTCTCCCGTG AAAATATCGT GTCTCTGTCC AGTGCTGTCC AATCAATTTT | 2100 |
| <i>15</i> | GATCCAACTA TACGATTGTC ATACGTAATT AAACTGCCAT TTGCTTGTTG ATAAAAAAT | 2160 |
| | ATTTGACCAA TTAACGTGAT AGCTAACGGG AATAAAAATC CACATAATAC CATAGTTATT | 2220 |
| | ATCGTTAAAC AAATACTATT TCTTATCGTA TTCATGGTAC AGGCTCCTTC CTCTTTACAC | 2280 |
| io | AAAAAATTGT ACAATCATAT CTATTAATTT AATGCCTAAA AACGGGACGA TTAATCCACC | 2340 |
| | TAATCCATAA ATCAACATAT TATTTATAAA GATTCTATCA ATGCTGTAAC CCTTTACTTT | 2400 |
| | TACACCTTTC ATGGCAATTG GAATTAAGGC AACAATGATT AATGCATTGA ATATCAAAGC | 2460 |

| | AATTGTTGAC | ATCATTAGTG | CAGGTAAAAT | TGCAAAGTAT | TTTGCTACGT | CATTAGCCAA | 2580 |
|----|------------|------------|------------|------------|------------|------------|--------|
| 5 | ACTAAATGTC | GTTAATGCAC | CTCTCGTCAT | TAATAATTGT | TTGCCTATTT | TTACAACCTC | 2640 |
| 5 | TATTAACTTT | GTAGGATTCG | AATCTAAATC | AATTAGATTA | GCTGCCTCTT | TAGCACTAAT | 2700 |
| | TGTCCCTGAG | TTCATAGCTA | ATCCTATATT | CGCTTtGTGc | tAGCGCAGGT | GCATCATTTG | 2760 |
| 10 | TACCATCTCC | TGTCATCGCA | ACAATATGGC | CTTTCGCTTG | TTCATCTTTG | ATGACTTTAA | 2820 |
| | TTTTATCTTC | GGGTTTACAC | TCTGCAACAA | ATCTATCAAC | CCCGGCTTCT | TTTGCAATTG | 2880 |
| | TAGCTGCTGT | TAAAGCATTA | TCACCTGTAC | ACATAACTGT | TTCAATCCCC | ATTTTTCTCA | 2940 |
| 15 | ATTCAGTAAA | TCGTTCTACA | AGACCATCTT | TAATCACATC | TTTTAAATAA | ATCACGCCAA | 3000 |
| | GCATGACATT | GTTTTCAATG | ACTATTAAtG | GnGTGCCACC | TTTACTCGAT | ACATCCATAC | 3060 |
| | AGAGAGACTC | AATATTAAGA | GGAATATTGC | CTTGTTGTTG | TTTGACAAGA | TTTATCATAC | 3120 |
| 20 | TATTAGGTGC | ACCTTTGAAT | ACCGATATTT | CATTTGTAAT | GATTCCGCTC | ATTCTAGTTT | 3180 |
| | CAGCTGTAAA | AGGCTTATAT | GTGCCATCAA | TGTCTTTAGG | CAGCTCATTT | ATATACATCT | 3240 |
| 25 | GCTTCGCTAA | TCGTACAATA | CTTTTTCCTT | CTGGCGTATC | ATCGTAGATT | GATGACATAT | 3300 |
| 23 | AAGCAGCGAC | TATCAATTTT | TCAAGCATTT | GTTGATTCAC | TGGTAAAAAT | TCACTAGCGA | 3360 |
| | TTCGATTGCC | ATAAGTGATT | GTGCCTGTCT | TGTCTAAAAT | CATTACATCG | ACATCTCCAC | , 3420 |
| 30 | ATACTTCTAC | AGCACGCCCA | CTTTTCGCTA | ATACATTGAA | TTGAGTAACA | CGATCCATGC | 3480 |
| | CTGCAATACC | AATCGCCGAT | AACAAACCAC | CGATTGTCGT | TGGTATTAAA | CATACTGTTA | 3540 |
| | ACGCAATGAG | CATCGCAATA | GGTAAAATTA | AATGCAGGTA | AGATGCTATT | GGATATAACG | 3600 |
| 35 | TTACAATAAC | GACTAAAAAT | ATAATTGTTA | ACGTTGTTAA | TAATGTAAAA | AGTGCAATTT | 3660 |
| | CATTTGGTGT | TTTATTTCTT | TCCCCCCTT | CAACTAAGGC | AATCATTTTA | TCTAAAAAAG | 3720 |
| | ATGTACnCGC | TTCACTCTCA | ACACGTATTT | CTAACCAATC | AGATGTTACA | AGTGTACCGC | 3780 |
| 40 | CAATGACTCC | ATCAAAATCG | CCACCTGATT | CTTTTATCAC | AGGTGCAGAC | TCACCAGTAA | 3840 |
| | TTGCAGATTC | ATCAACGGTT | GCTAATCCAT | TTATTACAAC | GCCATCAGCA | GGGATTGTTT | 3900 |
| 45 | CTCCATTTTC | TACCCGAATA | TTTTGTCCGG | CTTTTAACTC | TGTGGCGTTC | ACTATCCGAT | 3960 |
| 40 | ACGCACCATT | TTCTTCTATC | AATCGAGCAG | TTAAATTTGA | TTGTGCTTGT | CTTAAACTAT | 4020 |
| | CAGCTTGCGC | TTTTCCACGA | CCTTCAGCAA | AGGCTTCTGA | AAAATTAGCA | AACAATATAG | 4080 |
| 50 | TTATTAATAA | TATGATAAAA | ATTGTAATCA | AATAACCTCG | CGATAGATAG | CTAGTTCCAA | 4140 |
| | ATATGTCAGG | AAAACATATT | AATATCAACG | TTAAAATCAT | TCCAACCTCA | ACGACAÄACA | 4200 |
| | TTATCGGATT | TTTTATTAAT | TGTTTAAGAT | TCAGCTTATA | AAAACTCATT | TTCAAAGCTT | 4260 |

737

| | TITATTITAA | AGTTAAAAAT | TCACCAATAG | GACCAAGTA | TAGTACTGGA | ATAAATGTCA | 4380 |
|------------|--------------|------------|--------------|------------|--------------|------------|-------|
| 5 | AACCACTTAG | TAAAACGATA | AATACGATTA | GTGATACGCC | AAAATAAGGT | TTATCAATCG | 4440 |
| 3 | CTATTGTATA | TTTATCTTGA | TGGTATGATT | TTTTATTCAC | TAAACTTGAT | GCAATCATTA | 4500 |
| | ATTGCAAAAT | AATTGGTATA | TAACGAGAAA | GCAACATAAT | GATTCCTGTA | GAGATATTCC | 4560 |
| 10 | AGAATGTTGT | ATCATCTTTC | AGTCCTTCAA | ACCCTGATCC | ATTGTTCGCA | GCAGCTGATG | 4620 |
| | TCATTTCATA | CATAACTTGT | GAAATACCAT | GAAAAGACGG | ATTCGTtATa | CTTtCACTTG | 4680 |
| | CTCCAGGAAT | CATAAAAGCA | AGTGCTGAAA | ATACTAAAAT | TAAAATTGGG | TGTATGAGAA | 4740 |
| 15 | AGACTAAGAC | AATACATTTC | ATTTCACGGG | CGCCAATTGG | CATATTTAAA | TATTCTGGTG | 4800 |
| | TTTTACCAAC | CATCAAACTG | CATATAAACA | CCGTCAGTAA | GACAAATATC | AATAAATTCA | 4860 |
| | TGAGTCCTAC | GCCTTCGCCA | CCAAATACAA | CATTTAGCAT | CATTAATACC | ATTGGTCCTA | 4920 |
| 20 | ATCCACCTAT | AGGCGTTAAG | CTATCATGCA | TGTTATTAAC | AGAACCCGTT | GTAAATGCCG | 4980 |
| | TCGTAATAAC | TGTAAATAGT | GCTGACAAAC | CTGCTCCAAA | CCGTACCTCT | TTACCTTCCA | 5040 |
| 25 | TATTCGGTCC | ATAAATGCCT | AAATTCGCTA | GTATTGGATT | ACCACGATAC | TCACTCCACA | 5100 |
| | TAGTTAATGT | AAGAATTGCT | ATAAAAATGA | AAAACATTGC | GACAAATAAT | ATCAACGCAT | 5160 |
| | GACGATGTAC | TCGTTTACCA | TGTCTACTTA | ACATGCGACC | AAATAAGAAC | AACATTGACA | 5220 |
| 3 <i>0</i> | TAGGAAGTAA | CATCATACTG | CCCATTTCTA | TAAAATTGCT | CCAAATATTT | GGATTTTCAA | 5280 |
| | AAGGTGTTGC | AGAATTTCCT | GCTAAAAATC | CTCCACCATT | CGTACCAAGA | TGTTTTATTG | 5340 |
| | ATTCAAGTGA | TGCAATAGGT | CCAAATGCAA | TATGTTGAAT | ATGTCCGCTT | AAAGTCCGAA | 5400 |
| 35 | TCATTAAATT | AGCATGCAAC | GTTTGTGGTA | Caccttgagt | CATCAATAAA | ATACTAATTA | 5460 |
| | AACATGATAA | rggtaaaagt | ACTCGGACAA | TAAACCGAAC | AATATCTTGA | TAAAAATTAC | 5520 |
| | CAATGATATT | AGTTAATCCA | GTTAAACGTC | TCAACATCGC | TATACAAACG | GCGTAACCTG | 5580 |
| 10 | ATGCACTAGA | IGTAAACATT | AAATATGTCA | TTACAATCAT | TTGCGTTAAA | TATGTCACAT | -5640 |
| | CTGaTTCACC (| GTTATAGTGT | TGLAAATTAC | TATTTGTTAA | AAAAGATATT | GCTGTATTAA | 5700 |
| 15 | ACGCTAAATC | TATCGATTGG | TTTAAATTAT | GATTTGGATT | TAAAAAAAGC | CATTGCTGAA | 5760 |
| | CTATTAGCAA 1 | TACAAATGTT | ATAAACCCCA | TAAATCCATT | AAATGCCAGA . | AAATGTTTGA | 5820 |
| | CATATGTTTT A | AGCTGACATG | TGTTCTAAAT | CTGTGCCGAT | AATTTTAAAA | CACATATTTT | 5880 |
| 50 | CAAATCTAGT A | AAATATTAAA | TCTACTCTTG | ACGATTGCAC | CAATGCTACG | CGATATAGAT | 5940 |
| | ATCCACTAAA A | ACATACGTA | ATCATAACCA | TCATTGTTAG | AAACAAAATT . | ATTTCCATGA | 6000 |
| | TAACCCTCAC 1 | TTATATATT | TCTAAAATTT : | TTCACTACGA | ATTAAGGCAT | AAAATAAATA | 6060 |

| ACACAACAAC ATCGTAACAA CTTGTTTATG AGAGAAATNT TAATTTTCAA ACTTAGTTAT | 6180 |
|---|------|
| TAAGAAANCA TTAAGATGTG TATGCAGAAA TAAATTTTAT AGCATTTAAT TGTGAAGAAT | 6240 |
| ATTATGATAT TGCTATCGAG GTGAAGGTTA TG | 6272 |
| (2) INFORMATION FOR SEO ID NO: 141: | |

2) Intoldation for SEQ ID No. 141

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1978 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

15

20

25

30

35

40

45

50

10

5 .

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 141:

AAATGATGTT TTACAATAAA TATANAAACG TATCAACATA TATCATCATA TTTTTAGTTT 60 CAAGTGCAGC CTTTGCAATA TTCTTGTTAA GTGCGnACAT TAGTGCTCAC TCGGAACAAG 120 TGTACGAAAT GACTGACCAT CAAATTAAGA ACAATACGAT AAATAAAGCA TACGAACATA 180 AAGACCCTAC AAACAATAGC GAACAAAGAG ATGGGAAAGT GTTCGCTTTA ATAAATTGAT 240 300 ACATTGTCAC AACGTTATTT TGCCTATTTT TGCGMAATAG CGTTTTTTAT TACWTTTTTG CTGATSTTAA ATTTGTTATA TTTTGTTAAA GTATTATAAT GATTGAATAA ACAAATTGAA 360 GGTAGGTTTT TTAATTGAGT AATTCTGATT TGAATATCGA AAGAATTAAC GAGTTAGCTA 420 AAAAGAAAA AGAAGTAGGA TTAACTCAAG AAGAAGCAAA GGAGCAAACA GCCTTAAGAA 480 AAGCTTATCT TGAGAGTTTT AGAAAAGGGT TTAAACAACA AATTGAAAAT ACTAAAGTAA 540 TTGATCCAGr AGGTAATGAT GTAACACCTG AAAAAATTAA AGAGATACAA CAAAAAAGAG 600 660 ATAATAAAA TTAAATCACA AATCTGTAAA GAATTTTCTG ACATTATAAC TTGAAATAAG TATTTACTT ATCTTTTAT TTTAAAATAA GTTATAATGT ATTTGATAAA ATTGAAGAAG 720 GGAAGATACA CAAGATGTTT AATGAAAAAG ATCAATTAGC TGTTGATACG CTACGTGCAC 780 TAAGTATCGA CACAATCGAA AAAGCGAATT CTGGTCATCC AGGATTACCT ATGGGAGCTG 840 CCCCAATGGC TTACACTTTG TGGACACGTC ATCTGAATTT TAATCCACAA TCTAAAGATT 900 960 ACTICAATAG AGACCGTTTC GTATTATCTG CAGGGCATGG TTCAGCATTA TTGTATAGCT TGTTACATGT TTCTGGTAGT TTAGAATTAG AAGAATTAAA GCAATTTAGA CAATGGGGTT 1020 CTAAAACACC AGGTCATCCT GAATACAGAC ATACAGATGG TGTAGAAGTT ACTACCGGAC 1080 CACTTGGACA AGGTTTTGCT ATGTCAGTAG GATTAGCTTT ACAGAAGATC ACCTAGCAGG 1140 GAAATTTAAT AAAGAAGGAT ATAATGTTGT AGATCATTAC ACATATGTAT TAGCTLCTGA 1200

55

כבי יהו טוטובט ו

| • | AAGTAAATTA | GTTGTTTTAT | ACGATTCAAA | TGATATTTCA | TTAGATGGCG | AATTAAACAA | 1320 |
|---|-------------|-------------|-------------|------------|------------|------------|------|
| | AGCTTTTTCT | GAAAACACAA | AAGCTCGTTT | TGAAGCATAT | GGTTGGAATT | ACTTACTAGT | 1380 |
| | TAAAGATGGT | AATGATTTAG | AAGAAATTGA | TAAAGCGATT | ACTACAGCTA | AATCTCAAGA | 1440 |
| | AGGACCAACG | ATTATTGAAG | TTAAAACAAC | AATCGGATTT | GGTTCACCGA | ATAAAGCAGG | 1500 |
| | AACTAATGGT | GTTCATGGGG | CACCTTTAGG | TGAAGTTGAA | AGAAAATTAA | CATTCGAAAA | 1560 |
| | TTACGGTTTA | GATCCTGAAA | AACGTTTTAA | TGTTTCAGAA | GAGGTATACG | AAATTTTCCA | 1620 |
| | AAATACTATG | TTAAAACGTG | CTAATGAAGA | TGAATCTCAA | TGGAATTCAT | TATTAGAAAA | 1680 |
| | ATATGCAGAA | ACATATCCTG | AATTAGCAGA | AGAATTTAAA | TTAGCGATTA | GTGGTAAATT | 1740 |
| | GCCTAAAAAT | TATAAGGATG | AATTACCACG | TTTTGAACTG | GGTCATAATG | GTGCATCTCG | 1800 |
| | TGCTGATTCT | GGTACTGTTA | TTCAAGCAAT | CAGTAAAACT | GTCCCTTCAT | TCTTTGGTGG | 1860 |
| | ATCAGCAGAC | CTTGCTGGTT | CAAACAAATC | CAATGTAAAT | GATGCAACTG | ATTATAGTTC | 1920 |
| | TGAAACACCT | GAAGGTAAAA | ATGTGTGGTT | TGGTGTACGT | GAATTTGCTA | TGGGTGCT | 1978 |
| | (2) INFORMA | TION FOR SE | Q ID NO: 14 | 2: | | | |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7588 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 142:

| TAGTAGTATT | TATTAAATTA | TACGAAGGGA | CCCAACACAG | AAAATTCATT | TTATTGAATT | 60 |
|----------------|------------|------------|------------|------------|------------|-----|
| TTACATTTAT | GTGCCAAGTT | GGGAAAAATG | TCTTATTTTT | TCaAAGTATT | TAAAAGTAAA | 120 |
| ATTACATGTT | AATACGTAGT | ATTAATGGCG | AGACTCCTGA | GGGAGCAGTG | CCAGTCGAAG | 180 |
| ACCGAGGCTG | AGACGGCACC | CTAGGAAAGC | GAAGCCATTC | AATACGAAGT | ATTGTATAAA | 240 |
| TAGAGAACAG | CAGTAAGATA | TTTTCTAATT | GAAAATTATC | TTACTGCTGT | TTTTTAGGGA | 300 |
| TTTATGTCCC | AACCTTTTTA | GAATATTAAA | TTTCTACAAT | TTCGTCATCT | TCAACAATAA | 360 |
| AGCCCATTGT | ATTGACGCTG | TTATTTAAGA | AAGTCAGAAT | ATAACGCATT | ACTTCATCAC | 420 |
| GTTCTGGCTC | ATTGTGAACC | TCGTGGTAAA | AACCTTGCCA | AGCTTTAAAA | TATAATTCAG | 480 |
| GTGTTTGATA | TTTTTCTTTA | AACTCATCAA | TTGCCCTAGT | ATCAACAATT | AAATCCTTCG | 540 |
| TTCCATACAT | TAATAGCGTT | GGCATTGGTT | GAATGTCATG | AATATGAGCC | ATCGTATCTT | 600 |
| TCATCGTCTC | ATTAATTGTA | TTATACCAAT | GATACGTTGC | TTTTTTTAAC | ATTAAACCAT | 660 |

| | CATTAAAACG | TGTGTCTTT | GAAATTTTAC | CTATATTIGA | AACAAGTTTA | TCTTTACGAT | 780 |
|----|------------|------------|------------|-------------|------------|------------|------|
| | TTTTTCCATT | CTTTTGAAGT | TCTAGCATAG | GAGAAATTAA | CATCATCCCC | TCGATTGGCA | 840 |
| 5 | ATTCTACTTT | TTCAAGTAAA | TTTAATAAAA | TCAAACCGCC | AAGTCCTACC | CCTAATACAT | 900 |
| | AAGTAGGAAT | TTTATATTCA | TTAGCTATCT | TTAACCAGTC | TAGCAAACTT | TCGTGATACG | 960 |
| | TTTGAAAGTT | TTCAATTTGT | CCTTTATTAG | CTCTTGAAGT | TTGACCTTGA | CCAGGCAAAT | 1020 |
| 0 | CTCCCATAAT | CACATGATAG | CCATTTCTTC | TTAACATCGT | AATAACATAT | GCATATCTTC | 1080 |
| | CCGTATGTTC | TAATATATTA | TGAGCAATAA | CAACGACGCC | TTTCGCATCA | TTTTCAGCTT | 1140 |
| 5 | CCCACTTCCA | CATTATTATA | CTGCCCCTTT | TTCATTAATC | TTCAATAACA | TAATTATAGC | 1200 |
| 3 | AAATTCACTA | TGTAGATTTC | TATTTATAGT | ATTATTGTTG | TCCATATTAT | TATATATAAA | 1260 |
| | TGAAATCAAC | ATCAATAATA | GTGTAATTAT | ACATAATTAT | TTTTGATTGT | TTTTGATGAA | 1320 |
| 0 | AACGCTTTCT | CGAATATTTT | TTTCATGCTA | AACTTATTGT | AAACACAAGG | GTTTGGAGGA | 1380 |
| | GTAGCAATGG | CACTATTAAA | GAATTTTTTT | ATCGGATTAT | CTAATAATAG | TTTTTTAAAC | 1440 |
| | AACGCAGCAA | AAAAAGTGGG | CCCACGTTTG | GGCGCCAATA | AAGTCGTTGC | CGGAAATACA | 1500 |
| 5 | ATTCCAGAGT | TAATTAATAC | AATCGAATAC | TTAAATGACA | AGAATATCGC | TGTTACGGTA | 1560 |
| | GACAATTTAG | GGGAATTTGT | CGGTACAGTT | GAAGAAAGTA | ATCATGCTAA | AGAACAAATT | 1620 |
| | TTAACAATTA | TGGACGCGCT | TCATCAACAT | GGCGTAAAGG | CACATATGTC | TGTTAAATTG | 1680 |
| 0 | AGTCAGTTAG | GTGCAGAATT | CGACTTAGAA | TTAGCTTACC | AAAATTTAAG | AGAGATTTTA | 1740 |
| | CTTAAAGCAA | ATACTTACAA | CAATATGCAT | ATAAATATTG | ATACTGAAAA | ATATGCTAGC | 1800 |
| | CTGCAACAAA | TTGTTCAAGT | TTTAGATCGC | TTAAAAGGCG | AATTTAGAAA | TGTTGGTACT | 1860 |
| 15 | GTAATTCAAG | CATATTTATA | CGATAGCCAC | GAATTAGTTG | ATAAGTACCA | AGATTTACGA | 1920 |
| | TTACETTTGG | TTAAAGGTGC | ATATAAAGAA | AACGAATCAA | TTGCATTTCA | ATCTAAGGAA | 1980 |
| 10 | GACGTAGATG | CAAATTACAT | CAAAATAATT | GAACAACGTT | TGTTAAACGC | ACGCAATTTC | 2040 |
| | ACTTCAATTG | CAACACATGA | CCATCGCATC | ATTAATCATG | TAAAACAATT | TATGAAAGAA | 2100 |
| | AATCACATTG | AAAAAGATCG | TATGGAATTC | CAAATGCTCT | ATGGTTTTAG | ATCAGAGTTA | 2160 |
| 15 | GCAGAAGAAA | TCGCAAATGA | AGGCTATAAT | TTCACTATTT | ATGTACCTTA | TGGCGATGAT | 2220 |
| | TGGTTTGCGT | ATTTTATGAG | AAGATTAGCA | GAACGCCCAC | AAAACCTATC | TCTTGCTGTA | 2280 |
| | AAAGAATTTG | TGAAACCTGC | TGGCTTAAAA | CGTGTTGGCA | TAATTGCAGC | TTTAGGAGCT | 234 |
| 50 | ACAGTTATGT | TAGGTTTAAG | TACAATTAAA | AAATTATGCC | GTAAATAGAG | CAAGACATAA | 240 |
| | асаатааттт | AGGAGTCTCC | 33C33T33TC | A ATCTTCTAG | CCTCCTAAAT | СТТАТАТТСС | 246 |

| | TAGATTTTA | A TAAATTAGC | ATTTCAATTC | CACTTACTG | TGCTTCAGC | A CCTTTATTGC | 2580 |
|----|------------|--------------|------------|------------|--------------|--------------|------|
| | CAGCTTTCG | r Accigning | TCCACAGCT | GTTCAATACT | TTCAGTCGT | r aaaataccaa | 2640 |
| 5 | ATATGACTG | G TACATTAGTT | TGATCATTCA | CTTTAGAAA | ACCTTTCGC | ACTTCATTAC | 2700 |
| • | AAACATAAT | C ATAATGAGAC | GTAGCACCGC | GAATTACGCA | TCCTAATGT | ATTACTGCAT | 2760 |
| | CATAATTTC | TGATGAGGCT | ÄATTTTTTAG | CTACTAAAGG | AATTTCAAAC | GCACCTGGCA | 2820 |
| 10 | CAAATGCTAC | ATCAATATTG | TCTTCATTAA | CATCATGTCG | AATCAAAGTA | TCTTTTGCAC | 2880 |
| | CTTCAAGTAA | TCTTCCAGTG | ATAAAATCAT | TAAATCGACT | · AACTACGATI | GCAACTITCA | 2940 |
| 15 | AATCTTTTCC | ATTAATTTA | CCTTCAAAAT | TCATGTTAAA | ATCCTCCTAT | ATTAAATGAC | 3000 |
| | CCATTTTTAT | TTTTTTCGTT | TCCATATAAT | CATGATTATG | TACCGTTTCT | GGTACGATAA | 3060 |
| | CTTCAATTCT | TTCTGCAATA | TCAATGCCAT | ATTGTTTTAA | TCCCTCAAAT | TTACTTGGAT | 3120 |
| 20 | TATTACTTAA | TAAATTGATA | TGTTCGATGT | TAAAATATTT | TAAAATCTGT | GCAGCAATAT | 3180 |
| , | GATAATCTCG | CAAATCTTCA | TCAAAACCTA | ATGCTAAATT | TGCAGTTACT | GTATCATATC | 3240 |
| | CTTGCTCAAT | TAATTCATAT | GCGCGTAATT | TGTTTAACAA | TCCTATGCCA | CGACCTTCTT | 3300 |
| 25 | GAGGTAGATA | AATAATCATG | CCACCATGTT | CATTGATATA | CTTCATAGAC | GATTCAAGTT | 3360 |
| | GAGCACCACA | ATCACAACGT | TGACTATGGA | AAATATCGCC | TGTAAGGCAC | GCAGAATGTA | 3420 |
| | AGCGTACATT | TTCATGTTGT | CGAATTGCAC | CTTTTGTCAG | TACAACTATC | TCTTCATCTG | 3480 |
| 30 | TGTATGTCGC | TTTAAAACCA | TACATATCAA | ATGTTCCGAA | ATCTGTAGGC | ATTTTCACTT | 3540 |
| | TTGCCTTAAA | TTCAATTTCT | GGTTCTAATT | TTTTACGATA | TTCAATTAAA | TCATCAATCG | 3600 |
| 35 | TAATCATCTT | TAATTGATGT | TTTTCTTTAA | ACTTTTGTAA | ATCTTGTCCT | TTCGCCATCG | 3660 |
| ,3 | TGCCGTCATC | ATTCATAATC | TCACAAATGA | CACCAGCGGG | CTTGGCACCA | GTAAGTTTAG | 3720 |
| | CTAÃATCAAC | AGCCGCTTCT | GTGTGTCCAT | TTCTAGCTAA | TACGCCTTTA | TCTTGTGCTA | 3780 |
| 0 | CTAATGGAAA | TAAATGACCA | GGACGATTAA | AATCTTTAGC | TTCACTACTA | GGATCAATGÀ. | 3840 |
| | GCTTTTTGGC | AGTCAATGTA | CGTTCATAAG | CACTAATTCC | TGTTGTTGTA | TCTACATGAT | 3900 |
| | CAATACTCAC | TGTAAATTGC | GTACCAAAGA | TGTCGGAGTT | ATCATCAACC | ATTTGTACCA | 3960 |
| 5 | AATCCAAACG | TTGTGCAATA | TCTTTAGACA | CTGGTGCGCA | TATTAATCCC | CtTGCTTCTT | 4020 |
| | TCGCCATAAA | ATTAATGGTA | TTATCGTTCA | TCCATTCAGT | AACCGCTACT | AAATCACCTT | 4080 |
| | CATTTTCACG | ATTCTCATCA | TCTACTACAA | TAATTGGTTC | TCCATTTTTT | AAAGCCATTA | 4140 |
| 0 | AAGCACTGTC | AATATTATCG | AATTGCATGC | TACCCCTCCt | AAAAACCAAA | TGCTCTTAAT | 4200 |
| | TTATCTACAG | ATAATTGGTC | TTTATCTTTA | TTTAAAATAT | TTTCAACATA | TTTAAACAAA | 4260 |

| | CTCGTTTCTG | GAATAAGATG | AATGTCAAAA | CTGTTATCAT | GCTTATCAAA | TACCGTTAGA | 4380 |
|----|------------|------------|------------|------------|------------|------------|------|
| | CTAACACCAT | CCACAGTAAT | AGACCCTTGC | TTAACTAACT | GATTATTAAT | ATGTTGGCTA | 4440 |
| 5 | CATTGAATCG | TAATAATTTT | TGCATTGGCT | GTTTCATTTA | TTTTTGAAAC | TGTTCCTAGT | 4500 |
| | TCATCTACAT | GACCGAGGAC | AAAATGTCCA | CCAAACCTAC | CGTTACCACT | CATGGCACGC | 4560 |
| _ | TCTAAATTTA | CTTCTGATTG | TCGCTTAACA | TCTGCTAAAT | AGGTTTTATT | TTCAGTGCCT | 4620 |
| 10 | TTAATTACTT | GAACAGTAAA | AGATGTCTGA | TTAAAATCAA | TCACTGTTAA | ACATGCACCA | 4680 |
| | TTAACACTGA | TGGAATCACC | AATATGCATA | TCTGCCGTAA | TCTTATGTGC | TTCAATTTCA | 4740 |
| 15 | ATCGTCCTGA | CTGATTGACG | AATTTGAACA | CTTTTAACGA | CACCTATTTC | TTCAACGATG | 4800 |
| | CCAGTAAACA | TGCATCATCA | CTTCTTTCGT | AAAGTTAATT | TAACATTTTG | ATTTAATAAC | 4860 |
| | TCGGAATGAA | CAATTTCAAA | TTGGTTCGCA | TCTGGTATCT | CAATCACATC | ATTTGTTTGA | 4920 |
| 20 | TAAAATTGAT | AATTTCCAGA | TCCGCCAATT | AATTTCGGGG | CATAATAGAG | AATAAATTCA | 4980 |
| | TCTATATAAT | TAGATTGGAG | AAATTCTGAA | GTAGTGGTTG | GACCTGCCTC | GACTAGCAAA | 5040 |
| | GTTCCAACTC | CTCTTTTATA | TAAATTGTGA | AGAATTGTTG | TTAAATCGCA | AGACTTCAAG | 5100 |
| ?5 | TAAATAATTT | CAATATGTGT | TTGATTGGTT | GTTAAATTTG | GATTTTCAGT | ATATATCCAA | 5160 |
| | ATTGGTGTTG | ATTCATCTTG | ATAAATTTGC | TGATTAAAAT | GAATATTCCC | AGACTTAGAC | 5220 |
| | AATATTACTT | TTATAGGGTT | TTTTCCATCT | TGAATACGTG | TAGTATATTG | TGGATCATCT | 5280 |
| 30 | AATTCAACTG | TACGTCTTCC | AGTTAACACT | GCGTCGTGTC | GATGTCTTAA | CTTATAGACA | 5340 |
| | TCTTGTTTAA | CCTCTTTGTT | AGTAATCCAT | TGACTTTGTC | CATTATCATT | CGCTTGTTTA | 5400 |
| 35 | CCATCTAAAC | TTGCAGATAC | TTTCACTGTA | ATTTGTGGCA | GTTGCTTTGC | TTTTGCTTTA | 5460 |
| - | AAAAAGTCTT | GGTATAATTG | TGATGCCCGT | TCATCATCAA | CGCATTCAAC | CTCAATACCG | 5520 |
| | TGAÇCCCGTA | ACGTCTCATC | ACCATGTGTG | TCTAACGAAT | TGTCTTTTGT | TGCGTATACT | 5580 |
| 10 | ACTITICA | TCTTACAATC | AATTATTTTG | TTAACACAGG | GTGGTGTTGA | ACCAAAATGA | 5640 |
| | CTACATGGCT | CTAACGTAAT | ATAAATCGTC | GCACCTTCAG | CATTTTGTTG | TGCCATATCA | 5700 |
| | AGTGCTTGAA | CCTCCGCATG | CTTGTCACCT | TTTCTCAAGT | GTGCACCAAT | ACCAACAATC | 5760 |
| 15 | CTACCTTCTT | TAACTACAAC | AGCGCCAACG | GGTGGATTAA | CACCTGTTTG | ACCTTGTACC | 5820 |
| | ATATTTGCAA | GTTGAATCGC | ATAATCCATA | AATTGACTCA | AATGATCACC | TCTATAAACA | 5880 |
| | AAAATCCTCA | CATCATGAAT | TAAGATGCAA | GGAGAAAAT | TTATCGTTAA | ATAAGCCTAT | 5940 |
| 50 | TTGTACACAT | TTTTACAAAT | ACGCTACATT | ATCTTTGTCG | ATAATTAACA | TTCTTTCTCC | 6000 |
| | CATCCAGACT | TTAACTGTCG | GCTCTAGAAT | CTCACTAGAT | CAGCCACTAA | TATGAAACAT | 6060 |

| | TTATATATG | AATTGTTATA | GATTATTTGA | GTACGTAGTA | TGTCAACTAC | ATTTAAAATG | - 6180 |
|----|------------|------------|------------|--------------|------------|------------|--------|
| | ATACTATATO | TTTTCTGAAA | AAACAATTAA | TGACGGTTTT | AATTTAATAT | AATCTGAGTA | 6240 |
| 5 | CTATAGGCAT | CTCATTGATA | TGATTCTTAC | TAACAGACAT | TAAAATCAAA | CCTTCAATTC | 6300 |
| | GTCTCTATAG | AGCGTTCTCT | TTATTATCTT | CTAGTTACAA | ATTATTGATT | GtCACtGCGC | 6360 |
| | TGTTGTTGCT | CATTCGATTC | TAAAGCATCA | TATAATTGAG | ATACTGTATG | CGCAACTTGT | 6420 |
| 10 | TCTACAATCA | TTTTCACACC | GTTTCGTAGT | TTATTAACAC | CGTTTGTCAT | TTGACCTATC | 6480 |
| | GCAATCATAT | TTGTTAATGT | TCCAAACCTT | GGACTAATAA | CTTGATTGGT | TTCCGGAATG | 6540 |
| 15 | ATTTGTATGC | CTCCCATTGG | GTGTGCTTGT | ACAATTTGTC | TATTTTCAAG | ATTTCTAATT | 6600 |
| | AATTGATCAT | CTTGATCCAA | TTCATTTAAA | TGACTTTTTG | CACCTGTCGC | GTTAATGACA | 6660 |
| | ACATTATATA | TGTCTACTGA | TTCTTGGTTT | TTGTATGAAA | AATAATACAA | CTTGCCATaC | 6720 |
| 20 | ATGTTCACAT | CTTCTAAATC | TTTTTTCAAA | ATTAAAGACT | TATTTTCTAT | TAATTCAATA | 6780 |
| | ATTAGTTCAG | CAGTTCTTGG | AGGCATTGGA | TTTGAATTTA | ATTGAATCAT | CTTTGAGTAT | 6840 |
| | TTTTGATTAA | ATTGATGTTG | GTCTTCAATA | CTTAAGCTAT | TCCATATCCA | ATTTAAATTC | 6900 |
| 25 | TCTTTCAAAT | GTTCAATCAT | ACTTTGGAAA | ATGCCCaTTT | CTGTTGGACG | CGCTAAATCA | 6960 |
| | TACTTCAAAT | CTGCAATATG | ATTTCCTGTA | CGTCTATGTA | CTAATTTTTT | AAAATCAATG | 7020 |
| | TCATATTCAG | CACATTCTTT | ТАААААТААА | GAAACTAAAG | TATCAAGCGG | TGCATTGCCG | 7080 |
| 30 | AAATGATGTT | TTTTAATGTC | ATTTAATTTG | TCTTTAGTTA | AGTACTTGAA | TGTCACGTCT | 7140 |
| | ATCATTGTAC | CTCTTACACT | TGGTAAATGA | GCAGAACGAC | TCGTCATAGT | AATTGGTAAT | 7200 |
| 35 | TTTGGATGAT | GAGCAGCAAC | ATAACGGACA | ACATCTAAAC | TGGCAAGGCC | TGTACCAATA | 7260 |
| 55 | ATCGCAATAT | CGTCCAGTTC | ATTTACTTCG | TCTAACGTAT | TATATGTTGG | ATAAGGCGTA | 7320 |
| | gcgatatatc | CTTTTTTACC | CTTTAAGTTA | TATGGATCAT (| GGTAGGCAAA | TGTACCACAT | 7380 |
| 10 | | | | | | ACATATGTAA | 7440 |
| | TAAGTTAAAT | | | | | | 7500 |
| | TTAGTTGATA | TATTTGGATA | TTTTTTCGTG | AACATAGATA / | AATAAGATTT | CATATAATGT | 7560 |
| 15 | CCGAATACAA | ATCTCGGTAA | ATATGCAG | • | | • | 7588 |

(2) INFORMATION FOR SEQ ID NO: 143:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10320 base pairs (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 143:

| | nCTAGGTATT | TTAAACCTAA | TCTAGATAAA | CTAGCTTCGT | AAGCAGCTGC | TACATTTTCA | 60 |
|----|------------|------------|------------|------------|------------|------------|------|
| 5 | CGACCGAAAT | CCTCAAAATA | TAATTTTGAA | GTAATAAATA | AGTCTTCTCT | AGCAATACCA | 120 |
| | GTTGACTCCA | ATCCGGCACG | AATGCCAGCA | CCTACTTGTT | CTTCATTCCC | ATAAACTTTT | 180 |
| | GCGGTATCAA | TACTACGATA | TCCTTGTTCA | ATGGCATACT | TAACACTTTC | CATGCAATTT | 240 |
| 0 | TCATCATTTT | CCACACGAAA | TGTCCCTAAA | CCAATTTGTG | GCATCGTGTT | TCCATTATAA | 300 |
| | AATGTTTTAA | CCTCCATAAA | TATCGCCTCA | CCTTTTTGAT | GTATTATACC | CTGTTATCAT | 360 |
| 5 | AACAAATCTG | AGTTGAATAC | ATGAGAAAAA | ACACTTAGAG | CAATCAACCA | CTAAAATTCT | 420 |
| J | AGTAATATCT | CTCAAATATT | AATCAAATTG | TAAAAGTAAT | TCTGTTTAAT | TTATGACAAA | 480 |
| | CTAAAAAAGC | CGAAGTAACA | ACATATAGTC | ATCACTTCAG | CCTAACATTT | AATTGAATGA | 540 |
| 20 | TTCAATTTTA | TCCATCATTT | GTTGTAAGTC | TTCCACGTTG | TATTGAATAC | GACCATGGAA | 600 |
| | TACAAATTTG | TTAAAGAACT | CGTCTAATTG | TTCAGCACCG | ACAAGCACTT | TGACAGCACT | 660 |
| | ATTTTGATTA | TAATTTGAAA | TCGTTACATC | GCCTTCATTT | TTAAGATTAA | AGTATAAAAT | 720 |
| 25 | TGAAGTTGGT | GTATATTTGG | CACCTAATTC | TTTTTGTAAG | TCTTCAGCCA | ATTGTTTAAT | 780 |
| | CGCCTCAATT | TGATCTGAAT | AATTTACAAA | TGATAATGAA | CGTTTGTCAT | CATTTTGATC | 840 |
| | CATCACAATA | GTTTGCGGTC | TAGATTTATC | TAAATCCAAT | GTATCAAATA | CTTGTTCCAT | 900 |
| | TGGTGGTAAA | TCTTTAAATT | GACCGCCACT | AATACCATTA | TAAACATGAC | CTTTTAACAA | 960 |
| | TTGAGAATCA | АТААТАТААА | GACCAGTTCT | TGTTAATACT | AAATGACTAA | TTCGTTCAAT | 1020 |
| | ATTATTAAAG | CCATCCTTTG | GTAAAAAGAT | ATTTGCCATA | ATGTGCATAT | CTTCTGGTCG | 1080 |
| 35 | AATTCGTTTT | TCTTTAACTA | ATCTTTCACG | AATACCAATT | AATCTCATGT | CCGTTACATA | 1140 |
| | TTCĄCTATGA | TTTTTCGAGA | ACAATTTTAA | TGCGTCAATC | TCACGATCTT | TTGTACTAAC | 1200 |
| 10 | CATGTGATTA | TAATCTTCTT | GTTGTTTTGT | AATTGTCTTT | TTATTTTGAA | TACGCTCTTT | 1260 |
| | CTCTAAAGCT | TCTTCATGAG | ACTITITAAT | GTTTTGTTCT | TGTTGTTCAT | ACTITICITC | 1320 |
| | TGTTTGTCGC | TTAACTTTTT | TCTTACTACC | TAAGGCAACT | AAAAAAAGGA | CAAAAAAGAT | 1380 |
| 15 | TAATGCAATG | AgCTACTGCA | ATAATGAGTC | CAATGACTAT | CGGTGAAGAT | AAATCCATCA | 1440 |
| | CAACAACGCT | CCTTTTTAAT | ATATGAATAA | CTTTAATTAT | AATAGAaAAG | CTAAAGATTT | 1500 |
| | TCGATACATA | TTATCATTTA | TATACCGAAA | ATCTTTTATT | TAGCTATATT | CAATTCATCT | 1560 |
| 50 | TATTATTTTA | CTGCGTCTTT | TAATȚCTTCC | ACTITGTCTA | ATTTTTCCCA | TGGGAATAAG | 1620 |
| | ACATCTGTAC | GTCCAAAATG | ACCATAAGCA | GCAGTTTGTT | TGTAAATCGG | TTGTTTCAAA | 1686 |

| | AGTTGCCCTT | CAGAAACTTT | ACCTGTTCCA | AATGTATCAA | TTGCAATTGA | CACTGGTTCT | 1800 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | GCAACACCAA | TCGCATATGC | CAATTGTACT | TCACATTGAT | CTGCTAAACC | TGCTGCAACA | 1860 |
| 5 | ATATTTTTAG | CCACATAACG | TGCAGCGTAT | GCAGCTGAAC | GGTCTACTTT | TGTAGGATCC | 1920 |
| • | TTACCACTGA | AGCATCCGCC | ACCATGACGT | GCATAGCCAC | CGTACGTATC | AACAATGATT | 1980 |
| | TTACGTCCTG | TTAATCCTGC | ATCACCTTGA | GGTCCACCGA | TTACAAAGCG | TCCTGTAGGA | 2040 |
| 10 | TTGATGTAGA | ATTTAGTTTG | TTCATTAATC | AAGTTTTCTG | GAACAGTTGG | ATAAATGACA | 2100 |
| | TGTGCTTTAA | TGTCTTCTTG | AATTTGTTCA | AGTGTCACAT | CCTCAGCATG | TTGTGTTGAT | 2160 |
| 15 | ACGACAATCG | TATCAATACG | TACTGGGTTA | TCATTTTCAT | CATATTCAAC | AGTGACCTGA | 2220 |
| | ACTITACCGT | CTGGTCGTAA | ATAATTTAAC | GTACCATCTT | TACGCACATC | TGATAAACGT | 2280 |
| | TTTGCCAATT | GATGTGATAA | ATAAATTGCT | AGAGGCATAT | ACGTCTCTGT | TTCATTCGTT | 2340 |
| 20 | GCGTAACCAA | ACATTAAACC | TTGGTCACCT | GCACCTGTTG | CTTCAATTTC | TTCTTCGCTA | 2400 |
| | TCTTTATCAC | GATACTCTAA | TGCTTTATCC | ACGCCTTGTG | CAATGTCAGG | TGATTGTTCA | 2460 |
| | TCAATCGCAG | TTAAAATTGC | CATTGTTTCA | TAATCATAAC | CATATTTTGC | TCTTGTGTAT | 2520 |
| 25 | CCAATTTCTT | TAATTGTTTC | TCTAACAACT | TTCGGAATAT | CAACATATGT | TGTTGTAGAA | 2580 |
| | ATTTCGCCGG | CGATCAATGC | CATACCTGTT | GTAACAGTTG | TTtCACAAGC | TACACGTGCA | 2640 |
| 20 | TTTGGATCGT | CTTTTAAAAT | AGCATCTAAT | ATTGCATCTG | ACACTTGGTC | AGCGATTTTA | 2700 |
| 30 | TCTGGGTGTC | CTTCTGTAAC | AGACTCTGAA | GTAAATAATC | GTTTGTTATT | TAACATAGTT | 2760 |
| | TGCTCCTTTA | AATTTATATT | ACGAAAATTC | TCTCTCTGTG | AGCTAAATAA | AAAAGACCTT | 2820 |
| 35 | CTAACTATTA | ATATAGAGAG | AAGGCCTAAT | ACGTCCATTC | GCTCTTATCG | TTCAGACCTA | 2880 |
| | TTTGTCTGCA | AAcGGTTTGG | CACCTTTCTT | TTATAAAAAA | GAGGTTGCTG | GGTTTCATTG | 2940 |
| | GGTCCATGTC | CCTCCACCAC | TCAGGATAAG | AGAATCCGTT | AAAAATAATA | GTACCTAATT | 3000 |
| 40 | AATGAATTAA | TGTCAATTIT | TCACAAATAA | ATTTACAGTA | AAATATTGTA | GATTAATTAT | 3060 |
| | GTTAATGTGT | TATACTAATT | AAATGTAAAG | GCTTACATTT | AAATTATCGC | TTTGGAGGGA | 3120 |
| | TTTAGGATGT | CAGTAGACAC | ATACACTGAA | ACAACTAAAA | TTGACAAATT | ACTGAAAAA | 3180 |
| 45 | CCAACGTCAC | ATTTTCAACT | TTCGACGACA | CAACTTTATA | ATAAAATCTT | AGACAATAAC | 3240 |
| | GAAGGGGTAT | TAACAGAACT | TGGTGCTGTT | AATGCAAGTA | CTGGAAAATA | TACTGGTCGT | 3300 |
| | TCGCCTAAAG | ACAAATTTTT | TGTCTCTGAA | CCTTCATATA | GAGATAACAT | TGATTGGGGA | 3360 |
| 5 0 | GAAATTAATC | AACCTATCGA | TGAAGAAACT | TTCTTGAAGT | TATACCATAA | AGTACTAGAC | 3420 |
| | TATTTAGATA | AAAAAGATGA | ACTATACGTA | TTTAAAqGcT | ACGCTGGTAG | CGATAAAGAT | 3480 |

| | ATGTTTATTA | GACCTGAATC | AAAAGAAGAA | GCTACAAAGA | TTAAACCTAA | CTTCACTATC | 3600 |
|-----------|------------|---|------------|------------|-------------|------------|------|
| | GTTTCTGCAC | CACATTTTAA | AGCAGATCCA | GAAGTTGATG | GTACTAAATC | TGAAACCTTT | 3660 |
| 5 | GTCATTATTT | CATTTAAACA | CAAAGTCATT | TTAATCGGCG | GTACTGAATA | CGCTGGTGAA | 3720 |
| | ATGAAAAAAG | GTATCTTCTC | TGTAATGAAT | TATCTCTTAC | CGATGCAAGA | TATTATGAGC | 3780 |
| | ATGCATTGCT | CAGCAAACGT | TGGTGAAAAA | GGCGATGTTG | CATTATTCTT | TGGTCTATCT | 3840 |
| 0 | GGCACTGGTA | AAACAACCTT | ATCGGCTGAC | CCACACCGTA | AACTAATCGG | TGATGATGAA | 3900 |
| | CACGGCTGGA | ATAAAAACGG | GGTCTTTAAT | ATCGAAGGTG | GCTGCTATGC | AAAAGCAATT | 3960 |
| 5 | AATCTTTCCA | AAGAAAAAGA | ACCACAGATT | TTTGACGCAA | TCAAATATGG | TGCAATTTTA | 4020 |
| • | GAGAACACTG | TAGTTGCAGA | AGATGGTTCA | GTGGACTTTG | AAGACAATCG | TTATACAGAA | 4080 |
| | AACACGCGTG | CCGCTTATCC | AATTAATCAC | ATTGACAATA | TTGTAGTACC | ATCTAAAGCA | 4140 |
| 20 | GCACATCCAA | ATACAATTAT | TTTCTTAACT | GCGGATGCAT | TTGGTGTTAT | TCCACCGATT | 4200 |
| | TCAAAGTTAA | ATAAAGACCA | AGCAATGTAT | CATTTCTTGA | GTGGTTTCAC | TTCTAAATTA | 4260 |
| | GCTGGTACAa | GCGTGGTGTG | ACAGAACCTG | AACCATCATT | CTCAACATGT. | TTCGGAGCAC | 4320 |
| 25 | CGTTCTTCCC | GTTACACCCT | ACTGTTTACG | CTGATCTATT | AGGTGAACTT | ATCGATTTAC | 4380 |
| | ATGATGTTGA | TGTTTATCTT | GTTAATACTG | GATGGACTGG | CGGAAAATAT | GGTGTAGGAC | 4440 |
| | GTAGAATCAG | CTTACATTAC | ACACGTCAAA | TGGTAAACCA | AGCGATTTCT | GGCAAATTGA | 4500 |
| 30 | AAAATGCAGA | ATATACAAAA | GATAGTACGT | TTGGTTTAAG | CATTCCTGTA | GAAATTGAAG | 4560 |
| | ATGTACCGAA | AACAATTTTA | AATCCAATTA | ATGCTTGGAG | CGACAAAGAG | AAATATAAAG | 4620 |
| | CACAAGCAGA | AGATTTAATT | CAACGTTTTG | AAAAGAACTT | CGAAAAATTT | GGTGAAAAG | 4680 |
| 35 | TTGAACATAT | TGCTGAAAAA | GGTAGCTTCA | ACAAATAAAT | TTGAATACTA | AATCAAAACC | 4740 |
| | ACCCGTGTGA | ACGGGTGGTT | TGTTCTGCGG | CTATAAGCCT | TCCTTACTGG | CCAGCCCTAA | 4800 |
| 40 | AAGĠGCACTG | ACAAGTCAGC | CAACTGCACT | ACTATTCCAG | CAACCCTAAA | GGGTTACTCT | 4860 |
| | TITTTCTTTC | TTTTTTTATT | TTTCTCTCCA | GTGAAAGGAT | CTAAATATTC | TTCCATTGAG | 4920 |
| | ATTTGGTCTG | CAACGATATC | CTCTTGTAAT | TGATTACGAA | TATAATTTTC | AATCACTTTT | 4980 |
| 45 | TTATTTCTAC | CTACTGTATC | CACATAAAAT | CCTTTACACC | AAAACTTTCT | ATTTCCATAT | 5040 |
| | CTATACTTTA | AGTTAGCATG | TCTATCAAAT | ATCATTAAAC | TACTTTTTCC | TTTTAAATAG | 510 |
| | CCAACAAATG | ATGATACCCC | AAGTTTGGGT | GGTATACTAA | CTAACATATG | GATATGATCT | 516 |
| 50 | TTACATGCCT | CTGCTTCAAT | TATCTCTACA | CCTTTTCTTT | CACATAATTG | ACGCAATATA | 522 |
| | ATCCCTATAT | CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT | TTTTCCATAT | ATCACTTGTC | TTCTGTATTT | AGGTGCAAAG | 528 |

| | | | | | | _ | |
|----|------------|------------|------------|------------|------------|------------|------|
| | AAATAGCATC | TCCTCGTGTT | GATTATTTTG | GTTGGCTGAC | CAATATTTAT | TCTAGCACGT | 5400 |
| | AGAGATGCAT | TTTTTGTGAC | AATGGTAGAA | CCTTTTCtGa | ACCATACGCA | TAGCGTATGG | 5460 |
| 5 | TTTTCTTTTT | ACAATTAAAG | AGCCAACCGT | TGTTATAGTC | TAACAATGGT | TGGCTCCTCT | 5520 |
| | TATTTTATGT | GCTAAAAATT | TATAGGCAAT | TTTATTACAA | CAATGTACAT | TTAAGGTGAC | 5580 |
| | CTTCATGCCA | AAATCGCATC | ACTCATTTAA | TGGAAGCAGC | ACGTCTTCAT | ATAAAGTACC | 5640 |
| 10 | GATCCCTAAT | TCAACGCATG | TAGTACCACA | TCTTCAAAGC | TTGATAGTTC | CCATGCGCAC | 5700 |
| • | ACCACGTTTC | ATACTAGCTA | TGCGACTCAA | CTTGGTTCAT | AAACTCTTTA | ATATAAGTCA | 5760 |
| 15 | ATGTTTCAAC | CATCGCTGGT | GGTCTTGGCA | CATGTCCTTC | TGCCATTTGA | TAAAATGTTT | 5820 |
| | CATGCGTGGC | ACCTTTTAAC | TCTAGTTGGT | CCGCTAAATA | ATACGCATGA | TGAATACCAA | 5880 |
| • | CTTGCTGGTC | TTTCCCTCCA | TGTACAATTA | ATATTGGCGG | ACTGTTTTCA | TTAATGTTTG | 5940 |
| 20 | GAATCGCTTG | GCGTGCCTCA | TATGCCGCTC | GATCTTTTTT | CGGATGACCA | ATCATTCTTC | 6000 |
| | GTAGCATGCC | TCTTAAATCG | ACACGTTCTT | CATACATTAA | ATCAATATCT | GAGACACCAC | 6060 |
| | CCCAGATTGT | ATAACTTGTT | ACTGGTAAGT | CTTGAAATGT | CAACAATCCT | TGTAAACCAC | 6120 |
| 25 | CTCGCGAAAA | ACCAACCATG | TGGATAAATG | CATGTGGATA | TTTATCATGT | AGCAACCTTA | 6180 |
| | ATAATTGCGT | CACATCATTT | AAATCGCCAC | GGTAAAATTC | GTCTTTGCCT | TCACTCCCAT | 6240 |
| | TGTTACCTCG | GTAGTATGGC | CCAATCACTA | AAGTTTGACT | ATCTGAAAAT | TGCATTAATC | 6300 |
| 30 | TACCTGCGCG | CACACGTCCT | ACTTGACCTT | TGCCACCTCG | CAAATAAACT | ACAATGCGAT | 6360 |
| | TTACTTCATG | ATGTGGTGTC | ATCATTAAAG | CTTTTACTTG | TAAGTCATCT | GACAAATATG | 6420 |
| 35 | TAATTTCTTC | GAATTGATGC | GTAAAATATT | CAATTGGCAT | TCGTTTACGT | TTGATAAAAC | 6480 |
| | CCAAGTGATT | GCACCCTCTC | TACGCATTTT | AAAATGGTAC | TATCTTGCAG | TAAGAAACTC | 6540 |
| | CGTTGTGCGA | GTTCAATATC | ATTGATACAG | TTAAACAACA | CTGGCCCTGC | TGTTTCTAAA | 6600 |
| 40 | TAATCGTTCT | | | | | TACAAAATCA | 6660 |
| | GTTTGATCAT | | GGTATATTGT | GCTATGTAAT | | AACTTTGGCG | 6720 |
| | CCTGTTTCTT | CATATAATTC | aCGTGTAACT | GCTTCAGCAC | TACTTTCCCC | GCGTTCCCTT | 6780 |
| 45 | TTACCACCAG | GAAATTCAAT | CCCCGTAAA | TTATGTTTGG | TAAAAAGCAA | TTGATTTTTA | 6840 |
| | AACGTTGGAA | TAGCTAGCAC | ATGATTGCCA | TCTGCTATCT | CATTATCCTT | TTTAAATGTC | 6900 |
| | AAATTAACTT | GACGATTATC | TTTATCCCTA | AACTTCACGC | GCATCACATC | CCTACATTGT | 6960 |
| 50 | ATGTTAATAT | AATAGTTAAT | TACTATCGTT | GGAGGCATTA | ATTATGAAAA | AGATATTCTT | 7020 |
| | GGCGATGATT | CATTTTTATC | AACGTTTCAT | TTCGCCACTC | ACTCCACCAA | CTTGTCGTTT | 7080 |

| | CCTTTATTTA | GGTATCCGTC | GTATTTTAAA | ATGTCATCCG | CTTCATAAAG | GCGGCTTTGA | 7200 |
|-----------|------------|------------|------------|------------|------------|------------|------|
| 5 | CCCTGTTCCG | TTAAAAAAAG | ACAAGTCAGC | AAGCAAGCAT | TCACATAAAC | ATAACCATTA | 7260 |
| | ATATGGTTGT | AATTGAGTTA | TATCCACTAA | AGGGGGGCGA | AATTCGAGTC | GCCCCTCTTT | 7320 |
| | TAATATGCCT | GAATGCGCCA | CCACATCTTG | TTCAAAATAA | TAACCTGCTG | GTGTAACATC | 7380 |
| | TCCTGGATAA | TCACCTTTAC | GAGCAAGCAT | CGCTGTAAAA | TAGCGGCTTA | AACCATATTC | 7440 |
| 10 | GTACATGCCG | CCAATAACCA | CTTTTGCACC | ATGACTTTTC | AAAGTATCAA | TTGCCGTTTG | 7500 |
| | CACTTTATCA | ATGCCACCTA | GACGAAATGG | TTTTAATACA | ACAACTTTCA | CATTGTATAA | 7560 |
| 15 | TTCTATCAAA | TTAATTATGT | CCAACAACGA | TGTTGCCTTT | TCATCAAGGG | CTATTGGAGG | 7620 |
| | TATTGTTCCA | TCCGCTACTT | CATCAAGCAT | GGAGATATCT | TTAAATGGCT | CTTCGATATA | 7680 |
| | AAGAACCTGT | TCACGCGCTA | ATAACTGTAA | CTGTGTGAAA | TCTTGACGAT | CCAAGGACTC | 7740 |
| 20 | ATTTGCATCT | ATAACCAATT | GAAAGTGAAA | GTCTAATTCC | CGTAACACTC | TAATTTGATG | 7800 |
| | CATGATTTGA | GGCGTCCATT | TTAATTTAAT | TCTGGTCGGC | TTTGTTGCTT | TTAATGACTC | 7860 |
| | TAGTTGTTTA | TTTGATAAGC | CGCTCGCTGT | CGCTCCATAT | GCTACTGAAA | ATGAAGGCAG | 7920 |
| 25 | TACATGAAAC | ATTTGATACA | ATGCCATGAC | AATAGTTGCC | CTTGCAGCAG | GCGTATTTTC | 7980 |
| | CAATGAATCT | ACTAATTTTA | GTGCTGCTTC | ATACGTTTCA | AATGATTTAT | TTCTATTATC | 8040 |
| | TTCGAACCAT | TGCTCAATTA | CATGTTTCAC | TGAGGCAATT | GTTTCATGAT | CATACCAATC | 8100 |
| 30 | TGTTTGAAAA | GCGTTACATT | CCCCGAAATA | TGCATTTCCT | TTGTCATCAA | TCAATTCGAT | 8160 |
| | AAACAAACAA | TCACGATGCG | TTAAAGTGAC | TTTCGGTGTT | ACAATTTGTG | ACTTAAATGG | 8220 |
| 35 | CTCACTATAT | TTATAAAAAT | GCAAAGCTGT | CAACTTCATC | AAATCATCCT | CTATACAACT | 8280 |
| 55 | TATTTCTTTG | TAATTTACCT | GTTGATGTAT | AAGGTAAAGT | ATCAACCTTT | TCAAAGTGTT | 8340 |
| | TCGCTACTTT | ATATTTCGCT | AAATGTTGTG | ATAAATATGC | AATCAATTGT | GCCTTTGAAA | 8400 |
| 40 | TGTCACTTTC | ACTGACAAAA | TATAATTTAG | GCACTTGGCC | CCAAGTATCA | TCAGGATGCC | 8460 |
| | CTACACATAC | TGCGTCACTG | ATACCTGGAA | ATTGCLTCGC | TACCGTTTCA | ATTTGATATG | 8520 |
| | GATAAATATT | TTCACCGCCA | CTAATAATTA | AATCTTTACG | TCGGTCATAA | ATCATGACAT | 8586 |
| 45 | AACCTTCATG | ATCTATTTCA | GCAATGTCAC | CCGTATTAAA | ATAACCATTT | TCAAACGTAC | 864 |
| | CCGTTAAATC | TGTTGGATAC | AAATATACAT | TCATCACATT | GGCGCCTTTA | ATCATTAATT | 870 |
| 50 | CTCCATGACC | TTCTTTATTA | GGATTTTTAA | TTTTTACGTC | AACATTGGCA | CTTGGCATCC | 876 |
| | CTACAGTGTC | AGGACGTGCA | TGCAACATTT | CCGGTGTTGC | TGTTAAAAAT | TGCGAACATG | 882 |
| | TCTCAGTCAT | ACCAAATGAA | TTATAAATTG | GCAGGTTATA | TTGTAATGCC | GTCTCTATCA | 888 |

| | AACCIIGIIG | CATAAGCCAA | TTTAAAGTTT | GTGGCACAAG | CGAAATGTGC | GTGATTCGTT | 9000 |
|----|------------|------------|------------|------------|------------|------------|-------|
| _ | CATTTTTAAT | CATCGTTAAA | ATTTGTTCGG | CATTGAATTT | ATCAACAATG | CGCACAGTAA | 9060 |
| 5 | AACCTTCAAT | AACAGCTCTT | AAAAGTACAC | TGAGACCCGA | AATATGATAA | ATCGGCAAGA | 9120 |
| | CAGATAGCCA | ATTAGTGTCA | CGATCAAATC | CCAAGCTCTC | TTTACATCCG | ATTGCACTGG | 9180 |
| 10 | CATAATGATT | ACGAAACGTT | TGTGGCACCG | CTTTTTGAGG | GCCCGTTGTC | CCTGATGTAA | 9240 |
| ,, | ACATAATCGA | TGCAATGTCA | TCTAAATTAA | ATGATGTATT | TAATATGTTG | GACGGCGACT | 9300 |
| | CTTTCGGCAC | CACAGTTTCA | TTCGATGTTT | CATATTGGAT | ACCCATTGTG | TTGTCCAACA | 9360 |
| 15 | AACTGTTCGT | TGTAATATCC | CTTCCAGCGA | ATTCAATATC | ATCCAGCGAT | ACAATTTGAA | 9420 |
| | ACCCTCGTAA | TTCCAGTGGC | AAGGTACAAA | AAATCAATTG | TACATCGATT | GACTTCATCT | 9480 |
| | GATTCGTCAT | CTCATTAGGT | GTCAACCTTG | TATTAATCAT | CGCAATTTCA | ATATTTGCCA | 9540 |
| 20 | ACCAACATGC | ATGTATTAAA | ATGATCGATT | GAATCGAATT | ATCTATGTAT | AGCCCAACAC | 9600 |
| | GAGATTGTTG | ATAAGCCTTG | AGTCTTTTAG | CCAATAGACT | CGCTTCACAG | TATAAATTTT | 9660 |
| | GATAAGTATA | AGATTCTTGA | CCGTCTGTTA | TCGCAATATG | ATGTCCATTT | TGTTGTGCTT | 9720 |
| 25 | GTTTATATAA | CCAAAAGTCC | ATGCGTTATT | CCTCCAAAAT | CATTTACATT | ATAATTATAA | 9780 |
| | CGATTTTATG | ACATTCTAGC | AGTGGTTATG | TTTAAAAATA | TAAAAAAGTA | GACGAATTGA | 9840 |
| - | TGCATTGATA | TGATTGTTAT | AATGCTCAAT | ACATATCGTT | ATATCATTCG | TCTACTATTA | 9900 |
| 30 | TCAGTTATTT | TTATTTAATT | TTAGTGTCAT | TCTGTCATTT | TGATGTGGTG | ATTTACCCAT | 9960 |
| • | TGTTGCCACA | TCATCTGCAA | TGTCAATTGG | TATACGGTTC | ATGTCTTGTA | ATGCACTTAA | 10020 |
| 35 | ATGGAATACT | TCATCATCTA | AATTTTCAAT | GAGATATACA | TAATATGTTA | CCTTGTCCTT | 10080 |
| | TTTATATTTT | AACGTTTTCC | AAAAGTCCGG | CTTGCAATTC | AATACATTAT | CCGGAATATA | 10140 |
| | TTCĀĀTĀĀĀ | AAGTAACGTT | TGCTGCCTAC | TTTGTCTATG | AAATATTTTG | CAGTGCCTTT | 10200 |
| 40 | TTCTATACCT | CTTATATGTG | CATAGTCTGC | TGAAAAGTAA | ATACTACCTA | TTGTTTCATT | 10260 |
| | ATGTTGTTGT | ATTTCAAATC | GTTGGCCTAC | TATTTTATTA | TTTGTGCTAC | nGGGGACTTA | 10320 |
| | | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 144:

(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1477 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 144:

55

45

| | GTGTGGATTG | GATTTTAAAA | TCACCCTCAT | AAATACTGTC | ATCAATATGA | TAAGTTACAA | 120 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | TTTCACCTAT | TATTAAATCA | GCCCCATCTA | ATACATCTCC | AAGCAATATC | ATTTGCGmTA | 180 |
| 5 | GTTTACATTC | GAATCTCATT | TTCGCATCTT | TAATTCCTGG | CGTCTTAATC | GTTGTAGATG | 240 |
| | TTAAAAGTGA | TAATTCTGTA | CGACTCAACT | CACTGTCACC | ATATGCTAAC | GGCGCTGCAG | 300 |
| • | TCTCATTAAT | ATCTTGAACA | TTATCTTCGT | CTGTAATATG | CACAACAAAG | TCTCCAGTCC | 360 |
| 0 | GTTCTATATT | TAATGCAGTA | TCTTTTCTCT | TACCTCCTGC | ACGTTGAACT | GCAATAGCAA | 420 |
| | TCATTGGCGG | ATGATTATTA | ACAATATTAA | AAAAGCTAAA | TGGTGCTGCA | TTTACTGATG | 480 |
| 5 | CATCTTGATT | TAATGTTGTA | ACAAAAGCTA | TAGGTCGTGG | AATAATTGAA | CCAATTAATA | . 540 |
| | ATTTATAGTT | TTCTCTAGCA | GTTAATGATT | GTGCATCAAA | CGTATACATA | ATACCTACCT | 600 |
| | CTTTTCTAAG | TATATCTAGG | TATTTCTCCG | ATTTTGGTTA | ATTTAAACAT | CTATTCTCCT | 660 |
| 0 | CTGAAAATCA | CTTGTATTTA | TTTAGCAAAT | CTTTTGAAAT | ATGACACATA | TGCATATCTT | 720 |
| | CTGGATATTT | TTCTAAATGT | TGCTGATGTT | CTTCAGCACT | ŢTTAATGTAG | TTAGACAGCG | 780 |
| | GTAAGACTTC | CACTGCAATT | TGATCTCTGT | CTTTACGTCG | TTCAATGAAC | TGACGCGCTT | 840 |
| 25 | CAATTAAGTG | GTCATCTACA | CAACTATATA | AACCCGTTCG | ATACTTTTGT | CCAATATCAT | 900 |
| | TTCCTTGTTG | ATTCACACTG | TAAGGATCAA | TGATTTCAAA | TAAATAATTC | ATAATGTCTG | 960 |
| | TAATTGTTAA | CATACGATCA | TCGAAATGAA | GTTTGACACA | TTCAGCATAA | CCATCATACG | 1020 |
| 30 | GACCGTCTAA | TTTAGAGCTT | CTTCCATTTG | CTCTTCCTGC | TTCTGTATGT | ATAATTCCAG | 1080 |
| | GTATTGTTGC | AAAAAATGCT | TCAACACCCC | ATAAACATCC | TCCTGCTACA | TAAACAACTG | 1140 |
| 35 | CCATATTTAC | ACCTCATCAT | CCTTTTTTAT | ATTTTTAACA | AGGTTATACC | ATTTAATACC | 1200 |
| | GCCATGACAT | GATTCTGATA | CACCTTCATT | ACGATACCCA | TATTTTTCAT | AAAATGAAAT | 1260 |
| | TAATGATTCT | CGACATGTTA | ACGTTACACC | ATGTCGATGA | TGATTCTTAG | CAAGAGTTTC | 1320 |
| 10 | AAAATAGTTT | AGTAAGCGAC | CTGCAATACC | CTGACCTTGA | TAATTTGGTG | CTACAACAAG | 1380 |
| | ACCTAACACA | CTAATATAGC | CACCTTCACT | ATTATTTGTG | GAGACATTTT | TAAATAAATC | 1440 |
| | ATCGCTAATG | TAACGCTCTT | TTATGACTGG | ACCGTTG | | | 1477 |
| | | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 145:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 3976 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

45

| | AGGTGATTAT | CCTAAAAATG | CTCATGAGGT | CGCTATTAAT | GATAAGTTAG | CTGCAGACAA | - 6 |
|----|------------|------------|------------|----------------|----------------|------------|------|
| | CATTAGAGTO | GGGGATAGAT | TACATTTTAA | AAATAATTCA | ACTAGTTATA | GAGTTTCTGG | 12 |
| 5 | TATTTTAAAC | GACACAATGT | ATGCGCATAG | TTCCATTGTG | CTATTGAACG | ATAACGGATT | 18 |
| | TAATGCATTG | AATAAGGTTA | ATACGGCATT | TTATCCAGTG | AAAAATTTAA | CACAACAACA | 24 |
| 10 | ACGTGATGAG | CTTAATAAAA | TAAATGACGT | TCAAGTTGTG | AGTGAAAAAG | ATTTAACAGG | 300 |
| 70 | TAATATTGCG | AGTTATCAAG | CAGAGCAAGC | ACCGTTAAAT | ATGATGATTG | TTAGTTTGTT | 360 |
| | TGCTATTACA | GCAATCGTTC | TAAGTGCATT | TTTCTATGTT | ATGACGATTC | AAAAAATATC | 420 |
| 15 | ACAAATTGGC | ATTTTGAAAG | CAATTGGTAT | TAAGACAAGA | CATTTATTGA | GTGCGTTAGT | 480 |
| | TTTACAAATT | TTAACACTAA | CAATAATTGG | GGTAGGTATT | GCTGTGATCA | TCATAGTAGG | 540 |
| | ACTATCATTT | ATGATGCCGG | TAACGATGCC | TTTTTACTTA | ACAACGCAAA | ATATTTTATT | 600 |
| 20 | AATGGTGGGG | ATATTTATAT | TAGTAGCGAT | TTTAGGTGCC | TCACTATCAT | TTATCAAATT | 660 |
| | ATTTAAAGTG | GATCCTATCG | AAGCAATTGG | AGGTGCAGAA | TAATGGCATT | AGTCGTTGAA | 720 |
| | GATATCGTCA | AAAATTTCGG | AGAAGGTTTG | TCTGAAACAA | AAGTTTTAAA | AGGTATTAAT | 780 |
| 25 | TTTGAAGTGG | AACAAGGGGA | ATTTGTCATT | TTAAATGGTG | CCTCTGGTTC | TGGGAAAACA | 840 |
| | ACATTGCTAA | CGATATTAGG | CGGATTGTTA | AGTCAAACGA | GTGGTACAGT | GCTTTACAAT | 900 |
| | GATGCGCCAT | TGTTTGATAA | ACAGCATCGT | CCTAGTGATT | TACGATTGGA | AGATATTGGT | 960 |
| 30 | TTTATTTTTC | AATCTTCACA | TTTAGTTCCT | TATTTAAAAG | TGATAGAGCA | ATTGACACTC | 1020 |
| | GTAGGTCAAG | AAGCGGGAAT | GACCAAACAA | CAAAGTTCAA | CAAGAGCAAT | ACAACTTTTG | 1080 |
| 35 | AAAAATATTG | GTTTAGAAGA | TCGCTTGAAT | GTATATCCGC | ATCAGTTATC | TGGCGGTGAA | 1140 |
| | AAGCAACGTG | TTGCGATTAT | GAGAGCATTT | ATGAATAATC | CGAAAATCAT | TTTAGCAGAT | 1200 |
| | GAGÇCCACAG | CAAGTTTAGA | TGCCGATAGA | GCAACAAAAG | TTGTTGAGAT | GATACGTCAA | 1260 |
| 40 | CAAATTAAAG | AACAACAAAT | GATTGGTATT | ATGATTACAC | ACGATCGAAG | ATTATTTGAA | 1320 |
| | TATGCAGATC | GAGTGATTGA | ATTAGAAGAT | GGCAAAATAA | CTGATTAGTG | GCTTGTAAAG | 1380 |
| | ACGCTAAATG | TTAATGATTT | AAGACATAGT | AGTATAAAAG | TTAGATAACA | GAATACGATT | 1440 |
| 45 | TGGGTTTACA | AAAAACAGGC | TGGGACATTA | AGTTCTTAGG | CAATGTAAAA | AAGCTGATTT | 1500 |
| | CTATTAATTA | TTTGATAGAA | ATCAGCTTTT | TTGATATGTA | TTTTATAATG | TACAGCTCGT | 1560 |
| | TGCATTCATA | TAGCTTGAAG | TCACGTTTAA | AACCATATCT | ATCATTATGG | TATGCATATC | 1620 |
| 50 | TTTTAAAACC | TATTCTTTTG | TTATTAGGAC | ATATAAATTC | ATCATTAAGT | TCGTCATATT | 1680 |
| | TCCAATTTTG | AGTGTTAAAA | ATCTCACTOT | та а а ститист | ACTEMENT MOVED | TT | 7740 |

55 ·

| | CACTATCATA | ACATGCATCA | GCTACAATAT | ACTCCGGTAA | ATAACCGAAG | nTATTTTgAA | 1860 |
|----|------------|------------|-------------|------------|------------|------------|------|
| | TCATTGTTAA | AAATGGAATT | AAAGTTCTAG | TATCTGTTGG | GTTTTGAAAT | AGGTCATAGG | 1920 |
| 5 | ATAAAACAAA | TTGAGAATTT | GTCGCTATTT | GTAAATTGTA | TCCTGGCTTA | AGTTGGCCAA | 1980 |
| | AGTGTCTTAT | TTTTTTAAAG | TATTTAAAAG | TAAAATTACA | TGTTAATACG | TAGTATTAAT | 2040 |
| | GGCGAGACTC | CTGAGGGAGC | AGTGCCAGTC | GAAGaCAGGG | GCCCCAACAC | AGAArcTGAC | 2100 |
| 0 | ATATAGTCAG | CTTACAACAA | TGTGCCGGTT | GGGGTGGCTG | AGACGGCACC | CTAGGAAGGG | 2160 |
| | ACCCGTCATC | AAAAATTCTA | TTTATAGAAT | TTTACAGTAA | TGTGCCAGAT | GGGCATAGCG | 2220 |
| 5 | AAgcCATTCA | ATACGAAGTA | TTGTATAAAT | AGAGAACAGC | AGTAAGATAT | TTTCTAATTG | 2280 |
| | AAAATTATTT | TACTGCTGTT | TTTTTTAGGG | ATTAATGTCC | CAGACTCTTT | AGTTTATTTA | 2340 |
| | TTTTCAATAT | AACAATTGTC | TAATCAAGGA | TTAACGAATA | TTTAAAGATA | GTTTGACGCA | 2400 |
| 0 | ATATTAGAAA | CAACCTATAA | TAATAGTTTG | TTTGTGGATT | AACTATTATA | AATAAAAGCG | 2460 |
| | GCGTAAAGAC | ATATAAACCA | ACTACTTGAA | CAATATAACG | TTAATAACAA | TCTATACTGA | 2520 |
| | TACATTACGC | CTAGATAATC | TTTGATGAGC | ACATGTAAGA | AAAAGTGATA | TGGTGTATGA | 2580 |
| 5 | CTTCCGACAC | CATCGATAGA | TAAACCTAAT | TTTTGGGCTA | GTCGTAAGGC | GCGCAATACA | 2640 |
| | TGAAACTGAC | TTGTtACACA | AACAATTTTA | ACTGCTTCAT | GATACAAATT | GTTGATGATT | 2700 |
| | TGTTTAGAAT | ATAAAAAGTT | TGTGTATGTA | TTTATAGAGT | GAGATTCCAT | TAGTATATCT | 2760 |
| 0 | GTTTTATCAA | CACCATGTGC | AATCAAATAA | CGTTGCATAG | CTAAAGCTTC | AGAAATTGGT | 2820 |
| | TCGTCTGGTC | CTTGTCCGCC | AGATACAATG | ATCTTTGTTG | CTGATGCTTG | TTGTTGATAG | 2880 |
| 5 | ATATCAAGTG | CACGATCTAA | ACGCGCTGCA | AGCATTGGTG | TGACAAATTC | GGTAAAAATA | 2940 |
| | CCAGCACCTA | ACACAATTAT | GATATCAACT | TCTTTGTTGT | ATGATCTATG | TCTATATGAT | 3000 |
| | ACTGTCCAAA | CGAGATAACA | AATAAAGGTT | AGTAACAGGG | AAAGACATAA | TATAGCTAAC | 3060 |
| 0 | CACATAGACA | AACCTTTCAC | AATAGGTGAC | TGAATCGTAC | TTATAAATAG | AAGTGCTGAT | 3120 |
| | GTGTAGAGTA | CAAATTTATA | TGAAAAAGAT | AATAATTTT | TAATAAATAA | GCGACTAGAA | 3180 |
| | GTATGAGAAA | ATAAATATCT | ATGTTTGAAT | AGCATGATAA | TACTGATTAT | TATAAATGTT | 3240 |
| 5 | ACAAACATAG | ACCAAGGGAA | AGTATAGGTC | ATGATGCTAT | AGATGAGTGA | CAAAAATATC | 3300 |
| | GATATGACAA | CTAAGATGTA | GCATGTTAAA | TTTAACGTCA | GAGTATAGTT | GAAAATTAAC | 3360 |
| | GGACAAATAA | CGATAAGTAT | TAATTATTAAT | AATAAATTCA | ATAACATACT | GACACCTCGC | 3420 |
| io | TTATAATAAA | TATTAAATAT | AAATGTAGAT | GATTTAATTT | ATTAAAGCAA | GGAGAAAGCA | 3480 |
| | GCAACATGTA | AATCTTAATT | TGTTATATTA | TATATGGGTC | AATATTTTTG | TGTTTTTAG | 3540 |

| TATGGTAAAA | CATTTACAAG | ACCATATTCA | ATTTTTAGAG | CAGTTTATAA | ATAACGTTAA | 3660 |
|------------|------------|------------|------------|------------|------------|------|
| CGCATTAACT | GCAAAAATGT | TGAAAGATTT | ACAAAATGAA | TATGAAATTT | CATTAGAGCA | 3720 |
| GTCTAACGTA | TTAGGTATGT | TAAATAAAGA | ACCTTTGACA | ATTAGTGAAA | TCACGCAAAG | 3780 |
| ACAAGGTGTA | AATAAGGCCG | CAGTAAGCCG | ACGAATTAAA | AAGTTAATCG | ATGCTTAATT | 3840 |
| AGTTAAGTTA | GATAAACCAA | ATTTAAATAT | TGATCAACGT | TTGAAATTCA | TAACCTTAAC | 3900 |
| TGACAAAGGT | Agagcatatt | TGAAAGAACG | TAATGCGATT | ATGACAGATA | TTGCGCAAGA | 3960 |
| TATTACTAAT | GATTTA | | | | | 3976 |

(2) INFORMATION FOR SEQ ID NO: 146:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3346 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 146:

| | GCTACCTAGG | CATTTAAGAG | ATCAAAAAAT | GTATGAATAT | GAACGTTATT | TTTATGAGCA | 60 |
|---|------------|------------|------------|------------|------------|------------|-----|
| | AGAACTTAAT | GGCGTTGATG | aAGGGGAAAT | TTTAAAGAAG | TTAAAAGACC | CACAAGATGT | 120 |
| | TGCAGCTGAA | ACAAAAGCTA | GAAGTGTTAT | TGATTATGCT | GAATCTAAAC | CAACATTTGA | 180 |
| | AAATATTTCA | AGAGCTGTTG | CTGCTTCATT | AAGTTTAGGC | ATTCTATCTA | TTTTTGTCAT | 240 |
| | CCTTATACCA | GTATCTATAG | TTGGATTATT | TGTATTAGCA | TTATTTTTAA | TATCACTTTT | 300 |
| | GCTGCTGTTT | TGTCCAATTA | TTTTATTAGC | ATCAGCAATA | TCCAGAGGAA | TTGTGGACTC | 360 |
| | AATTAGTAAT | GTATTTTTG | CCATATCATA | TTCAGGATTA | GGATTAGTAT | TTATCATTGT | 420 |
| , | CATATTTAAG | ATTTTAGAAT | ACATTTATCG | TTTAATCTTA | AAATATTTAC | TTTGGTATAT | 480 |
| • | TAAAACTGTC | AAAGGAAGCG | TTAGAAAATG | AAGAAATTCT | TTTTTATTGG | GCTTTTAGTG | 540 |
| | ITTGTTGTCT | TTTTTACAGC | AGCAACCATT | ATTTGGTTCA | GCTATGATAA | AAACAAATAT | 600 |
| (| GGTACTAAAC | AATATGATAA | AACATTCAAA | gACGATGCTT | TTGACAATGT | ATCTATAAAT | 660 |
| • | ITGGATAGTA | CAGAACTTCG | TATAAAACGG | GGGAATCAAT | TTAGAGTTAA | ATATGATGGT | 720 |
| (| GACAATGATA | TATTAATTAA | TATAGTAGAT | AAGACGTTGA | AGATTAGTGA | TAAAAGGTCT | 780 |
| 2 | AGACAAGAG | GATATGCAAT | TGATATGAAT | CCTTTTCATG | AGAATAAGAA | AACGTTAACG | 840 |
| 2 | ATTGAAATGC | CTGATAAAAT | GATTAAACGT | TTAAATCTAT | CATCTGGAGC | AGGAAGTGTT | 900 |
| 2 | AGAATCAGTG | ATGTTGATTT | AGAGAACACA | AGTATTCAAA | GCATTAACGG | TGAAGTAGTT | 960 |

| | AGTAAAAGTA | ACATTAAAAA | TAGCAATATT | AAAGTTGTTA | TTGGTACGCT | ACAAATCGAC | 1080 |
|----|------------|------------|------------|------------|------------|--------------|------|
| | AAGAGTCAAA | TTAAACAATC | CATATTTTTA | AACGATCATG | GTGACATTGA | ATTTAAAAAC | 1140 |
| 5 | ATGCCATCAA | AAGTAGATGC | AAAAGCTTCT | ACTAAACAAG | GAGATATTCG | TTTTAAGTAT | 1200 |
| | GATAGTAAAC | CTGAAGACAC | TATACTAAAG | CTAAATCCGG | GAACGGGTGA | TAGCGTAGTT | 1260 |
| | AAAAATAAAA | CATTTACTAA | TGGLAAAGTT | GGGAAAAGCG | ACAATGTTTT | AGAATTTTAT | 1320 |
| 10 | ACGATTGATG | GTAATATCAA | AGTTGAATAA | ATAAAGGATG | TAAGCACCGA | TATTAGGAAG | 1380 |
| | CATAATTTCT | CTAATATCGG | TGTTATTTAT | TTGTTGGCAA | AAGTTAAGTC | GGTATCTATA | 1440 |
| 5 | TTGCCAGTAA | AGTGAGTGAT | ATTAAGGTCT | TGACCATCTA | ACCATGATTT | GAAATCTATT | 1500 |
| | ATTTCTGGTG | GCGCATTTTC | TCCCAATGTA | AAATATGCAG | TTAATGTTTC | AGGTTGATAC | 1560 |
| | ATTGATGTAT | GGATGGTGCC | AGACCAGCTT | TTGAATAGTT | TACTGTAAAT | TTCATACTGA | 1620 |
| 20 | GGATTATTGA | ATAACTTAAA | TGCTGTAGTC | ATATCTAAAT | TATCATTAGT | TTGTGAAATG | 1680 |
| | GTACGCGCCA | GTCTTTCTTT | AGATTCTTTT | GTATAATTAC | GATTTTCATG | TGTTAATATT | 1740 |
| | TCAAAATGAT | TTGTACATAT | ATTATCATAA | CGAACATCTA | TTGATCTCGG | TGTCACTTCA | 1800 |
| 25 | ACAATTGCAT | GGTTCAATGA | TTTGTCCATC | AGTATGTAGC | TAAATGAGCT | TCTGTGTGGT | 1860 |
| | ATTTCTTTCA | ATAATTGGAT | TGCTTCTGTT | ACATTTCGGC | AATTTTCAAG | AATTAGACGA | 1920 |
| | CCAATCATAT | AACATACAAA | ACCATTTGCT | GGTTTCTTCC | GGTGCATAAA | GTTATAGCCC | 1980 |
| 30 | ATAGTTAATC | CTGACTCATT | CATACCATCC | ATTCTTCCAG | TTACCCTTGA | TACAGGACCA . | 2040 |
| | ATTTGAGCTA | AACCGCTATC | TGTAGGTTGA | TAAAGTAAGT | AGCGACCATC | ATAAGTTGCA | 2100 |
| 15 | GGGTGGTAAT | CATAATTTCT | AACCATGAAG | TCTTTGCCTT | GAAAGACCGT | GCAaCCACTT | 2160 |
| | TCTTTTAAAT | CGGTAAAACG | ATAATGTCCA | AAGTTTAAAA | TAATTTGGCG | TGTTGGCATT | 2220 |
| | TTGAGTATAC | TTTGTAGTCC | CATTAATTCT | TCCCATATTT | GAGGTGCGTA | TGTTTGGAAT | 2280 |
| 10 | ATTTGATAAG | TTTCATTTAC | ATCTATATCG | AAACGTGGGA | CaCnTTTTTT | CCATTCTTTT | 2340 |
| | TCTCGATTTT | TTAGAAGAGG | TGTTTGTTGA | AGCCATTTAC | CAGTTTTAAC | ACCTAACTCG | 2400 |
| | AAATGTGAAC | CTCTAAAAGT | CATGATATCT | GATGTCACTT | GTTGCATATC | ATCGGCCCCT | 2460 |
| 15 | TTCTTTTTAG | TTGTAATATA | TTGTAAATAA | ATAGTAATCG | TATGTATATT | GAATGTCATG | 2520 |
| | TTAAATAAAG | TTATATTTTA | CTAAATGAAA | TATAAAATTG | TTTGAGGTGA | TTTCTCGGTG | 2580 |
| | TATAAGACTT | ATCAATCAGT | TAAAACATAT | TTTTATAGAT | GGTGGGGATA | TTGAGTTAAA | 2640 |
| io | AACTTAAAAT | CATCTTATCA | TAAATATCAA | TCTTAAGTTA | GCATTCACGA | TAATAGTCAT | 2700 |
| | TGTTAACATT | AGCATATAAG | GTCATGTCAC | GTTGAAACAG | AGGTTCCTCG | GCATTTTTGA | 2760 |

| | TTATTTAATG ATTATTCTAT ATATGATAGT ATAATGAAAT GTAGATAGGT ATTTAATTTA | 2880 |
|------------|--|------|
| | ACAGAGGTGA AATTGAGATG TGGAATTTTA TTAAATGLGT GKTTAAATTC GTATTTAGCT | 2940 |
| 5 | TAGTTGCTAT TACAACATTA GTTGCTGGTG TTGGTGTAGT AGCATTTGCT TATATCTTTA | 3000 |
| | AAAAAGATIT TGAAGATATT GAAAGAAAAA CTAAAGAAAT TATTTCTGAT ATTGAAAGTA | 3060 |
| | AAAATAACTA ATAACATTTA GAGGCTGGGA CATAAATCCC TAAAAAACAG CAGTAAGATA | 3120 |
| 10 | ATTTTCAATT AGAAAATATC TTACTGCTGT TCTCTATTTN ATCAMTACTE CGTATTGAAT | 3180 |
| | GGCTTCGCTT TCCTAGGGTG CCGTCTCAGC CTTGGTCTTC GACTGGCACT GCTCCCTCAG | 3240 |
| 15 | GAGTCTCGCC ATTAATACTA CGTATTAACA TGTAATTTTA CTTTGGAAAT ACTTTTAAAA | 3300 |
| .5 | AATAAGACAC TTTGGCCCAA CTTGGCACAT AAATGTAAAA TTCAAT | 3346 |
| | (2) INFORMATION FOR SEQ ID NO: 147: | |
| 20 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2375 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 2 5 | | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 147: | |
| | GTTGAAGAAA GAAATATAAC AGTCAATTAT AATTATAACC TTGTTGAAAT CGACGGTGAC | |
| 30 | AAAAAAGTGG CTACATTCGA ACATATCAAA GCATACGATA GAAAAACAAT AAGTTATGAT | 60 |
| | ATGTTACATG TAACACCACC TATGGGTCCC TTAGATGTAG TAAAAGAAAG TACACTTTCA | 120 |
| | | 180 |
| 35 | GATAGTGAGG GTTGGGTAGA TGTTAACCCA ACCACATTAC AGCATAAAAG CTACTCTAAT | 240 |
| | GTATTTGCAC TTGGTGATGC TTCAAATGTA CCTACTTCAA AAACAGGCGC ACTATTCGTA | 300 |
| | AGCAAGCACC TATCGTCGCT AATAATTTAT TGCAAGTGAT GAATAATCAA ATGTTAACGC | 360 |
| 10 | ATCATTATGA TGGTTATACT TCATGCCCTA TTGTTACTGG ATATAATAGG TTAATACTTG | 420 |
| | CAGAGTTTGA TTATAATAAA AATACTAAAG AAACAATGCC GTTTAATCAG GCCAAAGAAC | 480 |
| | GTAGAAGTAT GTATATTT AAGAAAGATT TATTACCTAA AATGTATTGG TACGGCATGC | 540 |
| 5 | TAAAAGGATT AATATAATAA AGTACAGAAA ACAATAAATT TTTAATGAAA AATCTTTTAC | 600 |
| | TATAAAAGAT TAAGTATTTA AATGACGTGT CAGTGTTGTG TTTATATGTC GTGAATTTTT | 660 |
| | AGCTCTAAAT AGTATAAGAT TGAAAAAGTT GTTACTGTTT TAAATGATCA CGATGAAGTC | 720 |
| 0 | ATTCAATAAG AATGATTATG AAAATAGAAA CAGCAGTAAG ATATTTTCTA ATTGAAAATC | 780 |
| | ATCTCACTCC TCTTTTTTTA ACCTTTATAC CTCATCCTCT AAATTATAC | |

| | AGATATTCAA | ACCACGTGTA | CTCAAAATGA | TAGCTTGGTA | TGTACCTCCA | ATAGTAATTT | 960 |
|----|------------|------------|------------|------------|------------|------------|------|
| | CAATAACTTT | GTCTGTTGAA | CACTAAGAGC | AATTTTAATT | TCATAATGTG | TTGTAAACAT | 1020 |
| 5 | TTTTTTTGAT | TGGAGTTTTT | TTCTGAGTTA | AACGATATCC | TGATGTATTT | TTAATTTTGC | 1080 |
| | ACCATTTCCA | AAAGGATAAG | TGACATAAGT | AAAAAGGCAT | CATCGGGAGT | TATCCTATCA | 1140 |
| 10 | GGAAAACCAA | GATAATACCT | AAGTAGAAAG | TGTTCAATCC | GTGTTAAATT | GGGAAATATC | 1200 |
| 10 | ATCCATAAAC | TTTATTACTC | ATACTATAAT | TCAATTTTAA | CGTCTTCGTC | CATTTGGGCT | 1260 |
| | TCAAATTCAT | CGAGTAGTGC | TCGTGCTTCT | GCAATTGATT | GTGTGTTCAT | CAATTGATGT | 1320 |
| 15 | CGAAGTTCGC | TAGCGCCTCT | TATGCCACGC | ACATAGATTT | TAAAGAATCT | ACGCAArCTC | 1380 |
| | TTGAATTGTC | GTATTTCATC | TTTYTCATAT | TTGTTAAACA | ATGATAFATG | CAATCTCAAy | 1440 |
| | AFATCTAATA | GTTCYTTGCT | TGTGTGTTCG | CGTGGTTCTT | TTTCAAAAGT | GAATGGATTG | 1500 |
| 20 | TGGAAAATGC | CTCTACCAAT | CATGATGCCA | TCAATACCAT | ATTTTTCTGC | AAGTTCAAGT | 1560 |
| | CCTGTTTTTC | TATCGGGAAT | ATCATCGTTA | ATTGTTAACA | ATGTGTTTGG | TGCAATTTCG | 1620 |
| | TCACGTAAAT | TTTTAATAGC | TTCGATTAAT | TCCCAATGTG | CATCTACTTT | ACTCATGCGT | 1680 |
| 25 | TTGATAAAAA | CTTAAATAAT | ATTAATTCGG | TCATCAGTGG | CGTTAAATCT | TTTATCATTT | 1740 |
| | TTAGTTATAG | TTGATAAATT | TATATTTATA | AGCATATATG | GATATTTCAT | CAAAAATTTT | 1800 |
| 30 | TATTTATATA | AATCCGAACT | GCATACATAT | TTGTTTAAAT | AAGAGGTATT | ATTTTTCGGG | 1860 |
| 30 | AAATTGCTGT | CTGAGTTAAA | AGGATTAGTT | TTATAAAATG | AGTTGAACTA | TAGCCAAAAA | 1920 |
| | CGATTAAAAT | ACTGATAATC | CATTTTTGLA | TTATGTTAGG | GACTTTTTTA | CTTAATTTTA | 1980 |
| 35 | ACCCTATTGG | aGCmAATATA | ATACTCCCTA | TTATAAGGAA | TAAGGCGTCA | TATAAAGGGA | 2040 |
| | TATAACCTTG | AATAAGTTTG | ATGACAAAAG | CACCAATTGA | AGATATAAAA | GCAATTACTA | 2100 |
| | TACTATTAGC | GACTACAGTA | TTCATTGGTA | ATTTGAATAA | AACCAATAAT | ATAGGAATAA | 2160 |
| 40 | | | | | | AGGCCAATGA | 2220 |
| | TAACTAATAA | ATATTTATTA | AATGAAGACT | TTTCGGAACT | AGGTTECACT | TTAATAAACA | 2280 |
| | TTAATGTTAA | TGCAAGTAAA | GCAATAATGA | TATATACCGT | ATTTACAAAT | GTAGCATCAA | 2340 |
| 45 | ATAAATTTGC | TAGAAATGCA | CCTAACATAC | TCCCT | • | | 2375 |

(2) INFORMATION FOR SEQ ID NO: 148:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6115 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 148:

| | GAGGTTTCTA GACAAGCTTT TAATAACTTA CCAAACTCAT TAAGITGGTT GTGLTGGACT | |
|----|--|------|
| 5 | | 60 |
| | GCCLATTATC MAAGLATTAT GAGTTGTTTA ATATTAGLGC TAARACATAC GAAGAGTGGT | 120 |
| | TTAAACAATT TAGTAGTAAG AAAGCACAAT TCAGTATTAA TCTCACGGAT AAATGGATAA | 180 |
| 10 | TTCAAATCGC ATATGGTAAA TTAATAATAA TGGCTAAAAA TAATGGCGAT ACATATTTTA | 240 |
| | GAGTTCAAAC AATTAAAAAG CCAGGTAATT ATATTTTTAA CAAATATCGA TTAGAGATAC | 300 |
| | ATTCTAATTT ACCAAAATGT. TTATTTCCGC TTACAGTGAG AACACGACAA AGTGGCGATA | 360 |
| 15 | CATTTAAACT GAATGGGCGC GATGGTTATA AGAAAGTGAA TCGCCTGTTT ATAGATTGTA | 420 |
| | AAGTGCCACA GTGGGTTCGG GATCAAATGC CAATCGTATT GGATAAACAA CAGCGCATTA | 480 |
| | TTGCGGTAGG AGATTTATAT CAACAACAAA CAATAAAAA ATGGATTATA ATTAGTAAAA | 540 |
| 20 | ATGGAGATGA ATAGCGTTAT GCATAATGAT TTGAAAGAAG TATTGTTAAC TGAAGAAGAT | 600 |
| | ATTCAAAATA TCTGTAAGGA ATTGGGAGCA CAATTAACAA AGGATTATCA AGGTAAACCA | 660 |
| | TTAGTATGCG TGGGTATCTT AAAAGGCTCA GCAATGTTTA TGTCAGATTT AATTAAACGA | 720 |
| 25 | ATTGATACCC ATTTATCAAT TGATTTCATG GATGTTTCTA GTTATCACGG AGGCACTGAG | 780 |
| | TCAACTGGTG AAGTTCAAAT CATTAAAGAT TTAGGTTCTT CTATTGAAAA TAAAGACGTA | 840 |
| 22 | TTAATTATTG AAGATATCTT AGAGACTGGT ACTACACTTA AGTCAATTAC TGAATTATTA | 900 |
| 30 | CAATCTAGAA AAGTTAATTC ATTAGAAATA GTTACTTTAT TAGATAAACC AAACCGTCGT | 960 |
| | AAAGCGGACA TTGAAGCTAA GTATGTAGGT AAAAAAATAC CAGATGAATT TGTTGTTGGt | 1020 |
| 35 | TACGGTTTAG ATTATCGTGA ATTATACCGA AACTTACCAT ATATCGGTAC GTTAAAACCT | 1080 |
| | GAAGTGTATT CAAATTAATT TTTTAATCAA TTTCAGTTAT TATTACTATG CGTTTGAGAA | 1140 |
| | ATAATAGTGT AGACTCAAAA ATATGAAAAA TGTATTTCAT ATATATTTAA TTTTAGACAA | 1200 |
| 40 | GACATATGTC TTGAAAAGTT GAAAAATATA GAGATTGATA AAACTAATAC GGGTGTGAAT | 1260 |
| | GACATTGATG TTAAGCTCAA TTACTAGCTT ATAAAACATG TCATATGTTA CAATTTTTGT | 1320 |
| | TAGTTTTATT ATGGGAAGTA GGAGGAAATG ACGCATGCAG AAAGCTTTTC GCAATGTGCT | 1380 |
| 45 | AGTTATCGTA ATAATAGGCG TTATTATTTT TGGTCTATTT TCATATTTAA ACGGTAATGG | 1440 |
| | AAATATGCCG AAACAGCTTA CATATAATCA ATTTACTGAG AAGTTGGAAA AAGGTGACCT | 1500 |
| | TAAAACTTTA GAAATCCAAC CACAACAAAA TGTCTATATG GTAAGTGGTA AAACGAAAAA | 1560 |
| 50 | TGATGAAGAC TATTCATCAA CTATTTTATA TAACAACGAA AAAGAATTAC AAAAAATTAC | 1620 |
| | TGATGCTGCT AAAAAGCAAA ACGGTGTAAA ATTAACGATT AAAGAAGAAG AAAAACAAAG | 1680 |

| | TTTCTTCCTA | AGCCAAGCAC | AAGGTGGCGG | TAGTGGCGGT | CGTATGATGA | ACTTTGGTAA | 1800 |
|------------|------------|------------|------------|---------------|--------------|-------------|------|
| | ATCTAAAGCA | AAAATGTACG | ATAATAATA | ACGTCGTGTT | CGTFTCTCTG | ATGTAGCAGG | 1860 |
| 5 | GGCAGATGAA | GAAAAACAAG | AATTAATTGA | AATTGTTGAT | TTCTTGAAAG | АТААТАААА | 1920 |
| | ATTCAAAGAA | ATGGGATCTA | GGATTCCTAA | AGGTGTCTTA | CTTGTTGGAC | CTCCAGGTAC | 1980 |
| | TGGTAAAACA | TTACTTGCTA | GAGCGGTTGC | AGGTGAAGCT | GGCGCACCAT | TCTTCTCTAT | 2040 |
| 0 | TAGTGGTTCA | GACTTTGTAG | AGATGTTTGT | TGGTGTTGGT | GCGAGCCGTG | TTCGTGACTT | 2100 |
| | ATTCGATAAT | GCTAAGAAAA | ACGCGCCTTG | TATCATCTTT | ATCGATGAGA | TTGATGCTGT | 2160 |
| 5 | TGGTCGTCAA | CGTGGTGCAG | GTGTTGGTGG | CGGTCATGAT | GAACGTGAAC | AAACCCTAAA | 2220 |
| | CCAATTATTA | GTTGAAATGG | ATGGTTTCGG | TGAAAATGAA | GGTATCATTA | TGATAGCTGC | 2280 |
| | TACAAACCGT | CCTGATATCC | TTGACCCAGC | CTTATTACGT | CCAGGTCGTT | TTGATAGACA | 2340 |
| 20 | AATTCAAGTT | GGTCGTCCAG | ATGTGAAAGG | CCGTGAAGCA | ATTCTTCATG | TTCATGCTAA | 2400 |
| | AAACAAACCA | CTTGATGAAA | CGGTTGATTT | AAAAGCAATT | TCACAACGTA | CACCTGGTTT | 2460 |
| | CTCAGGTGCT | GATTTAGAGA | ACTTATTAAA | TGAAGCATCT | TTAATTGCTG | TACGTGAAGG | 2520 |
| 25 | TAAAAAGAAA | ATTGACATGA | GAGATATCGA | AGAGGCAACG | GATAGAGTTA | TAGCCGGACC | 2580 |
| | TGCTAAGAAA | TCTCGAGTTA | TTTCTAAGAA | AGAACGTAAT | ATTGTTGCTC | ATCACGAAGC | 2640 |
| | TGGTCATACA | ATTATCGGTA | TGGTACTTGA | TGAGGCAGAA | GTAGTGCATA | AAGTTACTAT | 2700 |
| 10 | TGTTCCACGT | GGACAAGCAG | GTGGTTATGC | AATGATGCTA | CCTAAACAAG | ATCGTTTCTT | 2760 |
| | AATGACTGAA | CAAGAGTTAT | TAGATAAAAT | CTGTGGTTTA | CTTGGTGGAC | GTGTATCAGA | 2820 |
| 35 | AGATATTAAC | TTTAACGAAG | TATCAACAGG | TGCTTCAAAT | GACTTCGAAC | GTGCAACACA | 2880 |
| ,3 | AATCGCACGC | TCAATGGTTA | CGCAATATGG | TATGAGTAAA | AAATTAGGAC | CATTACAGTT | 2940 |
| | CGGTCATAGC | AATGGTCAAG | TATTCTTAGG | TAAAGATATG | CAAGGTGAGC | CTAATTATTC | 3000 |
| 10 | AAGCCAAATC | GCATATGAAA | TTGATAAAGA | AGTTCAACGA | ATCGTTAAAG | AACAATACGA | 3060 |
| | ACGTTGTAAA | CAAATTTTAT | TAGAGCACAA | AGAACAATTA | ATTTTAATTG | CTGAAACATT | 3120 |
| | ATTAACAGAA | GAAACATTAG | TTGCTGAACA | AATTCAATCA | TTATTCTACG | AAGGTAAATT | 3180 |
| 1 5 | ACCTGAAATT | GATTATGATG | CAGCTAAAGT | TGTTAAAGAT | GAAGATTCTG | AATTTAATGA | 3240 |
| | TGGTAAATTC | GGTAAATCTT | ATGAAGAGAT | TCGTAAAGAG | CAATTAGAAG | ATGGACAACG | 3300 |
| | TGACGAAAGT | GAAGATCGTA | AAGAAGAAAA | AGATATTGCT | GAGGATAAAA | AAGAAGCTGA | 3360 |
| 50 | TAAATCTGAT | GAAAAAGATG | AACCAGCACA | TCGACAAGCC | CCAAATATCG | AAAAACCTTA | 3420 |
| | | | | ~1 mm1 m1 mm~ | 1 cm1 ccmcmm | TOTATO TA A | 2400 |

| | AATTGTTATA | GCAGAAAATA | ATTGTAAAAC | AAGTTACTTC | ATTATTTAGA | ATGATGGGTG | 3600 |
|---------|------------|-------------------|------------|--------------|---------------|-------------|------|
| · | TAGAATAAGT | ACAATTGTTG | CATTTTATGA | AGTAAAGTAA | TTTTTTAAAT | ATAGAGTAAT | 3660 |
| 5 | AGAGGAGATT | GAAATAATGA | CACACGATTA | TATTGTTAAA | GCATTAGCAT | TTGATGGAGA | 3720 |
| | GATTAGGGCT | TATGCTGCTT | TGACAACTGA | AACTGTTCAA | GAAGCACAAA | CGAGACATTA | 3780 |
| | TACATGGCCG | ACAGCATCTG | CTGCAATGGG | AAGAACAATG | caCAGCAACA | GCTATGATGG | 3840 |
| 10 | GCGCAATGTT | GAAAGGTGAT | САААААТТАА | CTGTCACTGT | AGATGGCCAA | GGACCTATTG | 3900 |
| | GACGAATTAT | TGCCGATGCA | AATGCTAAAG | GCGAGGTGCG | TGCTTATGTA | GACCATCCAC | 3960 |
| 15 | AAACTCATTT | TCCATTAAAT | GAGCAAGGTA | AACTTGATGT | AAGACGAGCG | GTAGGGACAA | 4020 |
| | ATGGATCTAT | TATGGTTGTT | AAAGACGTTG | GAATGAAAGA | CTATTTCtCT | GGAGCAAGTC | 4080 |
| | CaATTGTTTC | AGGAGAACTT | GGTGAAGATT | TTACTTATTA | TTATGCTACA | AGTGAACAAA | 4140 |
| 20 | CACCTTCATC | GGTAGGTCTT | GGTGTATTGG | TAAATCCTGA | TAATACGATT | AAAGCAGCAG | 4200 |
| | GAGGATTTAT | CATTCAAGTT | ATGCCAGGTG | CCAAAGATGA | AACAATTTCA | AAATTAGAAA | 4260 |
| | AAGCAATTAG | TGAAATGACA | CCAGTTTCTA | aattaattga | ACAAGGATTA | ACGCCAGAAG | 4320 |
| 25 | GATTACTAAA | CGAAATCTTA | GGTGAAGACC | ATGTGCAAAT | TTTAGAGAAA | ATGCCTGTTC | 4380 |
| • | AATTTGAATG | TAATTGTAGT | CATGAGAAAT | TTTTAAATGC | TATTAAAGGA | TTGGGCGAGG | 4440 |
| | CTGAGATTCA | AAATATGATT | AAAGAAGATC | ATGGTGCTGA | AGCAGTATGT | CATTTCTGTG | 4500 |
| 30 | GAAATAAATA | TAAATATACT | GAAGAAGAAT | TAAACGTGTT | GCTAGAAAGT | TTAGCGTAAT | 4560 |
| | TTAATTTAAA | TCAATACGCT | AAAATGTTTA | TTTTTAGCGG | TTTAGTGAAA | TGTAGAACTA | 4620 |
| 35 | AATAGTTGTA | TAATCCTTAG | TGATTTTGTT | TGCTTTCTAG | AATTTATTTG | TAATAAATA | 4680 |
| 33 | TCTATATCCG | ATAAATAAAC | TAAGATTTCA | ACAACTAACT | AAAAAGGAGT | GTTCTTAATG. | 4740 |
| | GCAĢĀAAAAC | CAGTAGATAA | TATTACTCAA | ATTATTGGCG | GTACACCGGT | AGTCAAATTG | 4800 |
| 40 | AGAAATGTAG | TAGATGACAA | TGCAGCAGAT | GTTTATGTAA | AATTGGAATA | TCAAAATCCA | 4860 |
| | GGTGGTTCTG | TAAAGGATAG | AATTGCTTTA | GCAATGATTG | AAAAAGCAGA- | GCGAGAAGGC | 4920 |
| | AAAATTAAAC | CTGGCGATAC | AATTGTAGAA | CCAACAAGTG | GTAATACAGG | TATCGGTTTA | 4980 |
| 45 | GCATTTGTAT | GTGCTGCTAA | AGGATATAAA | GCAGTATTTA | CTATGCCCGA | AACAATGAGC | 5040 |
| | CAAGAGCGTC | GTAATTTATT | AAAAGCATAC | GGTGCGGAAT | TAGTTTTAAC | GCCTGGATCA | 5100 |
| | GAAGCGATGA | AAGGTGCAAT | TAAAAAAGCT | AAAGAATTGA | AAGAAGAACA | TGGTTACTTC | 5160 |
| 50 - | GAGCCACAAC | AATTTGAAAA | CCCTGCGAAC | CCTGAAGTTC | ATGAGTTAAC | TACAGGTCCT | 5220 |
| | CACTTATTAC | 3 3 C 3 3 TTTTC 3 | 2000222200 | ATCCA TOCOCT | macons acomae | D0DD00000 | |

| | GTTGCTATAG | AGCCTGAGGC | TTCTCCAGTA | TTGAGCGGTG | GTGAGCCAGG | TCCACATAAA | 5400 |
|----|-------------|-------------|-------------|------------|------------|------------|------|
| | TTACAAGGTT | TAGGTGCTGG | ATTTATTCCA | GGCACTTTGA | ATACAGAAAT | CTATGACAGT | 5460 |
| 5 | ATTATTAAAG | TAGGAAATGA | TACAGCGATG | GAAATGTCTC | GTCGAGTTGC | TAAAGAGGAA | 5520 |
| | GGTATTTTAG | CAGGTATTTC | ATCAGGTGCT | GCGATTTATG | CTGCCATTCA | AAAAGCAAAA | 5580 |
| | GAATTAGGAA | AAGGTAAAAC | AGȚAGTAACA | GTATTGCCGA | GTAATGGTGA | ACGCTACTTA | 5640 |
| 10 | TCAACACCTT | TATATTCATT | CGATGACTAA | TTAATGTCAT | TTAAAAGAGT | GAGTTATCTT | 5700 |
| | TTTGAGATAA | CTTGCTCTTT | TTTTCTACCA | TGTATATTTT | TAAAAATATG | AGCGTTAAAT | 5760 |
| 15 | TAAACATTTT | TCTGATAAAA | ATATCCAGTG | AATGATAAGA | TAATAAACGT | ACATACTAAT | 5820 |
| | aactagtaaa | TAGCAGGAGT | AAATTTTATT | AGAGTTAAAC | AATACATAAT | TAAAGGGTGG | 5880 |
| | TTAACATGAC | TAAAACAAAA | ATTATGGGCA | TATTAAACGT | CACACCTGAT | TCATTCTCAG | 5940 |
| 20 | ATGGTGGAAA | ATTTAATAAT | GTTGAATCAG | CTATAAATAG | aGTGAAAGCC | ATGATAGATG | 6000 |
| | AAGGTGCTGA | CATTATAGAT | GTTGGAGGTG | TTTCAACĠAG | ACCCGGTCAT | GAAATGGTTT | 6060 |
| | CATTAGAAGA | TGAGATGAAC | AGAGTATTAC | CTGTTGTTGA | AGCTATTGTC | GGTTT | 6115 |
| 25 | (2) INFORMA | TION FOR SE | O ID NO. 14 | 19- | | | |

(2) INFORMATION FOR SEQ ID NO: 149:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10401 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 149:

| TAGATACTGG | GnTAAAcaTc | AAAAATAtyT | GCtTaTTCaC | GTGTTTAcGc | TCCCtCAAAC | 60 |
|------------|------------|------------|------------|------------|------------|-----|
| GCAĀCGTTAA | TTGCGTGTAA | TCATTTAGTG | TGAATTCAGA | CGCTTCTTCC | ATGACTATGT | 120 |
| CTGATATGCC | TTTTATCGAC | TTTATTTTCT | CTGGGTTATC | TAATCCTTTA | AACAAAAAA | 180 |
| CTGCGCCGTT | TGGCAATTCA | ACTITGTTAT | CAGTCTTATT | CCAAAGGCAC | ATGTCCCAAA | 240 |
| TACCAAAGTT | TATCAAACAA | TCTTTAACAT | CTTCGAACAA | ACTATCTTTA | ATTGTTGATT | 300 |
| GTACTTTTCT | AAGCCACAGT | ATACGCCTAG | GATATTTCCA | ATCTTGCAAT | GCTTTGAGTA | 360 |
| CAACTTTTTG | TATAACGCCG | TGAGACTTAC | CGCTCGAACC | TCCACCGTAA | TGKACTTCAG | 420 |
| TGAAGTLATC | GTAATTGGTT | AGTATTTCGA | ATATGTTTCT | ATTGAAAACA | TTAGACGGTT | 480 |
| TGTTAAAGTT | TAATTTAACT | TTCGTCATCG | TACTCACCAA | TATTAATCTC | AATATTCTTC | 540 |
| TGAGTAATTT | CTTTTTTATC | GATATACGCA | CCATGTACTT | TTAGTATGTG | GTCAATAGAT | 600 |

. ;

55

50

30

35

40

| | TITAAATGGT CATATTTCTT ACTGTAAGCC TCTTGAGGTT CTCCTCTAGC AATAGAAGCA | 72 |
|------|---|------|
| | GATAACGCTA AAGCTTCTGT AATACTCATT AAACGCTCTT CTTGTATCTG TTCTAATCGT | 78 |
| 5 | TCTTTAATAT ATTCCGAAAC ATTAACATTT CTTAACAATC GACTTGCTAA AGACTCTGCT | 84 |
| | GTTTTCTTAC TATAACCTGC TGTAATTGCT GCTTTTTTAC CATTACATCC ATTCATTATA | 90 |
| 10 | TATTCATCTG CGAATCTCTT TTGTTTTTCG TTCATTTCAT | 960 |
| 10 | ATACGCTTTT TAAAATTAAA AAAGGATTGG CTATAATCAG CCAACCCACA TAGATCCTTT | 1020 |
| | ATTCCTAATT GCGATAAGGG AAACGCAGTA CGATAGTCAA TATCCTACAC TATCATAATA | 1080 |
| 15 | TCTCATTTAA GGTATCAAAA ACTGCCACTT TACTGCCAAT TTCAGTCTTC CCCTAACTCT | 1140 |
| | TCCGCCAATC TAGATATGAT TTTTCTTTTG ATTCTATGAG CAGTTCTATC AGAAATGTGT | 1200 |
| | ATGTCAACAC AAACTTTCAC TAATTCCTTT TTATTAAAAT AATACTCTTG AATGAATTCG | 1260 |
| 20 | CGTTCTTTCC TGCTTGATGT GTTGATTATA CGTTCAATAG CGCTCTTAAA CTCAAGGATT | 1320 |
| | TTACCTCTTC GTATACTACA AAGATAATTA GTTACTGCCA TTTCTGTTTT CGATGTATTA | 1380 |
| | GACGGTACAA ACTCCCCGCC TATATTTGTA TCTGTTGGAA TCCACGGTGT CATTATTTCA | 1440 |
| 25 | CTTCTTAAAT CTTCAAGTTG TTTATGATAA TTAGGATAAT CACACAACTC ATCTTCTAAC | 1500 |
| | TTTCGAACTG TTGATAATTT TAATCCGTAT TTCTTTTTAG TCATGAATAC CCTCCGTACA | 1560 |
| | AATATGTTTA ATCTTCAAAG TGTCTCAATC TACTTCTTAA TATCTCTATC TCTCGCTCTT | 1620 |
| 30 | TAACTTTTAC ATCACCTTTT AACTGTTCCG CTTGTAACAT CACACCAAAC AATAAGATGA | 1680 |
| | CTAGTAATAT AATTGCTATG ATTAACCACA TCATCTACTC CGACACCTCC GCCCTCATCA | 1740 |
| 35 | AATCAGACTG ATCACTCAAC TTTGCGAAGT CACTTGGCGC CTCTACATCA TCATTAGCCG | 1800 |
| ,,, | TCATCATAAT ATATACTTGC TCAGTTACAT ACTTACCTAA CTCATACATC GCTAGTAAGA | 1860 |
| | ATAATAGTCT CAAAATTTCT TTAACCACCA CTAAACACCC CATGTTAATT TATCGATAAT | 1920 |
| 10 | TTGTATAGCT TGTTTTAATG CGTCTCTTTT TTCTTTGATA TCTCTATTAT CGCCATCTTC | 1980 |
| | ATCAGCTGAC ATTAACTCAC TGTCATATTC ATATAATAGT TCTGATATTT CATTACTAGC | 2040 |
| | TACTACTAAT AAGTTTTCAT CTACATCAAT CGTTACCGTT TTCTTTGGCA TCTCCATCTC | 2100 |
| 15 | TCCTTATCTT AACTTGTGCC TCGTATTTGC GCTCAGCTTC TTCTTTACTC TCTGCCTCAA | 2160 |
| | CAACTGTAAA CGTCTGATTA TCTCTAGCAG TAGTAAAATG TTCATGTGGT TGTCCTGTTG | 2220 |
| | AATCTTTGAA TGTTGTGACT AAGTATTGCG TCACTTCTTA TCACTCCTTT GAATGATTCT | 2280 |
| io · | AAGTTTTTCT ACGAATAAAA GTATTAGTAC AACACTCAAT GTAGCCAACA TATTTTTTTG | 2340 |
| | CTTTGCAAAA TCTACTATAA CGATTAAGAC TAATAACATT CCAATTCTGC ATGTAAATAA | 2400 |

| | TACAAGTATT | GGAACTAATG | TAATGATGTA | ACTCACTTCC | CCAAAACCTC | CTTGACTCGA | 2520 |
|----|-----------------|------------|------------|------------|------------|------------|------|
| | TCTAAGATGT | CTTTACACTC | CGCTACTTCC | GAAGCCTTTT | TCTCCACGTT | CTGAAACACT | 2580 |
| 5 | TTCGAATTCC | TCCACTTGCT | TTAGTTCAGG | TGTCCATATA | GGCACGATAA | CCAATTGAGC | 2640 |
| | TAGTTTGTCT | CCTTCGTTGA | TTTGATAAGT | TCCGTATTGT | CTTATGGCGT | CACTCAAATC | 2700 |
| | GATTTCTCCT | TTAATATCAA | AAACACCTGG | TGTGATATAA | CCATTCGATG | CAATAGCGTC | 2760 |
| 10 | ATTCTTGATA | TTAATCCCTA | AATTGCCGTG | ATATCCCGCG | TCTATCTTGC | CTGTTTCAAT | 2820 |
| | CACTAAATGC | GTTTTACTAC | TTACACCACT | ACGGCTAGTT | AATAGTCCGA | CATAGCCCTC | 2880 |
| 15 | TGGTATGCTT | ACAGCTACAT | CTGTTTTAAT | CACTGCCTTT | TCTTGTGGCT | CAAGTACGAC | 2940 |
| | AGTTTCAGCT | GAGAATATGT | CATAACCTGC | ATCCGTCTTA | TGATTTCGTT | CGGGCATTCT | 3000 |
| | AGCATTTTCT | GATAATAGCC | TTACTTGTAA | TGTGTTAGTC | ATTTTCCTGC | TCCTCCCTAG | 3060 |
| 20 | CTGTAGCAAA | CGCTATTCTC | AATTTCAATC | TTTCAACAAT | ATGAATTAGT | GCGGTATTGA | 3120 |
| | GGAATATTTC | AAATTCTTCA | ATGTTCTCAT | CTATAAAATC | AAGTATTTCT | TCCTCTTGTT | 3180 |
| | CACTGTCAAA | CTCGCTTAGT | ACATCCCAAA | TATTTATGTC | GCTTTTGCTC | GTTTCTAATA | 3240 |
| ?5 | CTCTTTTGAT | TATTTCTGAA | TTACTTTTAT | TACTCATTTT | CCTTGTTCCT | CCTCATATTT | 3300 |
| | ATAGACAACT | TGACCTGCCA | TAATCCCTAC | TGCTTCATCA | AGTTCAATAC | CTTCTTTAAC | 3360 |
| | TGAATGTTGA | ATAGCATTTG | TCATTCCCTC | AAGTATTTCA | TCAAACGCTT | GTGCTCTCTT | 3420 |
| 30 | ATACACGTCC | TCAATCTCTT | TTAGTAATCC | CTCTGTGTCA | TTACCGTTAT | ACGCACTAGC | 3480 |
| | ACTGATCACT | GATTGTTCAA | TTTGTTCGCG | GTTATTCATC | ATTTCCATCT | CCTCTAAAAT | 3540 |
| | AAAGTTAGTT | GCTTCTGCTC | CTCGTATTCC | AAACCATGTT | GCTTTATATA | TGTTTCGAGC | 3600 |
| 35 | TCTTCCGCTG | TATCAAATGT | CTTTTTCACG | CCTTGCCAAC | CTGGCACGAT | ATGCCCATGa | 3660 |
| | AAGTAATAAG | TGCCGTTCAC | TACATGGATA | TGTGCCACTC | GTTCGTTATC | CTGATACAGA | 3720 |
| 10 | TATCTCTTAG | ATCCGAAAAA | TTGGTTTAAG | TATTCTTTAC | ATGCGCTATC | GGTTTTAGGC | 3780 |
| | ATTTATGCTT | CCTGCCATTT | CTTAAACATT | TGGTTATAAG | TAGTATCAAA | CCAGTACGGA | 3840 |
| | TCACGTGAAT | GTTTTTGAGG | CACATTAAAC | AAATGTGGCT | TCTTCTTACG | TAGTTCAGCC | 3900 |
| 15 | TCTTTACGTC | GTTGCCTAGC | CATTTCACGC | TCTTTGCTCT | CTCGCTCCAT | GATTTTGGAT | 3960 |
| | AACACAATTT | CTTTATACTC | AGCTAAGCGC | ATACCATAAG | GTGCATGTAA | GGCTTCTAAC | 4020 |
| | AACGCCCAGC | CACCTCGTAC | TCTTTTTGCA | ACCATTCCTG | GAGTTAAACC | GTTCTTTTTT | 4080 |
| 50 | ATCAATTCAT | TTTCATGTTC | GGTAAATTTA | TATGGTTTAC | CGTTAATCTT | TACGATACTC | 4140 |
| | A THIT A THICAR | COMOMNANA | | mmmacm | | | |

| | ACATTTAAGT | TAACCATCTC | AGCTTTTCCG | TTTTTATATC | CACTAATAGT | TGATCTTGAT | 4320 |
|----------|------------|-------------|------------|------------|------------|------------|------|
| | ACGCCAGTTT | CATTGTGCAA | ATCTTGGACA | CTTACGTTAT | CTCTAGCCAT | GATTACCCTT | 4380 |
| 5 | AAATTAGTTG | CGAATACTEC | GTTCAACTTC | ATTTATTCCA | CCTCTATATA | TGCATGTCTT | 4440 |
| | ATTGTTATGT | TGTCATACTT | TAGTAATTCG | TCCGGATTGT | CATCTAAGCG | CTTTGCCAGC | 4500 |
| | GTATCTTTTT | CTTTATCCAC | ATCATCGTAA | TGCTGATATT | CAACTTCTGT | AGGTATTCTT | 4560 |
| 10 | ATATCAATCG | TTGCGTTTAT | ATATGCTTGT | TGTTGCATTA | GATCACTTCA | TTTCTCTTTT | 4620 |
| | TCTTTTACGT | CTGACTTTCA | CTAAGTCCTC | ATATACCATC | CATTCTTGAC | CTGTGTATTT | 4680 |
| 15 | AGGCGCTTTA | CATATCCACG | TTAAATTCAC | ATCTCTATAC | TGATATCTGA | ATATCTTCGC | 4740 |
| | TTTGATGTTG | GCAACTTCAG | TCGCCTTACC | TTTAACGTCT | ATAACTTCAA | CCAGTTTCCC | 4800 |
| | TTCCTTCCAC | AAAGAGAAAT | CGGCTATATA | CGTAATCGGT | CTTTGTTTCC | CGAATTTAGG | 4860 |
| 20 | TTGTAATTCA | AATTTCGGTT | GTATTTCGAT | ACGATCATAG | TTAGTGCCAT | TCATATTACT | 4920 |
| | TTCTAAATAT | TGGTAATATT | CGCACTCTAC | TTTGCTATCA | AATACAATTC | CTTTGTACTC | 4980 |
| | AACTTTCTTA | GCATTGTATT | TACTCATTGT | GCCACCTCTA | AAȚATCAAAT | ATCGTTGCTT | 5040 |
| 25 | GCAATCCTAG | CTCTTGCTCA' | TATAGAAGCC | CGTGAGCGCC | TTTGAATCGT | TTTAGGTCAC | 5100 |
| | TATCAGTCAT | AATTTTCTTT | TCGTCGCTGA | AATGGGCTCC | TGTGAGCGAA | TAAACTTCAT | 5160 |
| | TTACGTTGTC | TTTATACTTG | ATGACCTTAA | TATCTTCCGT | GCCATCTTCT | CGGTATAAGT | 5220 |
| 30 | AATATTTTTC | TTTCGGCATT | TTTTAACACT | CCTTAATGTG | TGTTTTCTTC | CAGTTGATTT | 5280 |
| | CATTCATGAT | TTTCTTTTCA | ACTCTGTCGT | AATCATCGAA | AGGCGATAAC | TCGTTATTGT | 5340 |
| | CCAACAATCT | ATTGACCGCC | CAACCAGTCT | CGATATATAC | ATTTGCTACA | ATCGGGTCGC | 5400 |
| 35 | TTTGCTTTGT | CTCTTCATAC | ATCGATTTCA | ATAAGCTTTT | GAATTGCATT | ATATTCATGT | 5460 |
| | GAAAĀACCTC | TGAGTCTTCT | TGTAATACTC | AAATTCAATT | ATTCCGGTTT | CGCCGTCTTT | 5520 |
| 40 | GTTTTTGGCT | ATGTTACATT | CAACAATAGA | TTTGCCAGTG | ATACTGTCAT | CTTCGTCACG | 5580 |
| | GTTATAATAA | TCATCACGGT | AAAGTAGCAT | CGCTAAACTC | GCATCTGCTT | CTATTCCGCC | 5640 |
| | TGATTCTTTC | ATGTCCGATA | GCATTGGTCT | TTTATCCTGT | CTAGACTCGA | CACCACGATT | 5700 |
| 45 | CAGTTGTGAA | AGTAGTACGA | TGATTGCGCC | TGTCTCGTTA | GCGATTATCT | TTAAGTCACG | 5760 |
| | TGATATCTTT | TCTACTGCTA | CACGTCTATC | AACTTTCGCA | TCAGTATCCA | TCAGTTGAAG | 5820 |
| | ATAATCTATA | AAAATAACTT | GTTGCCTGTC | TGAATGCCTC | ATTGETGCGC | TCGCACATCT | 5880 |
| 50 | TGCGGTGTGA | TATTACTTTT | ATCAGAAATA | TCGATGCCTA | ATTTCATGAT | TTTATCCATC | 5940 |
| | GCATTCGTTA | ACTITICITIA | GTCATCCGGC | GTTAAGTTCC | TGATTTCTTT | TATCTTTGTT | 6000 |

| | AGACTAAAGA | AAGATGTTTT | GTATCCATTT | TGTGCTATGT | TCAGCATCAT | GTTTAATGCA | 6120 |
|------------|------------|------------|-----------------|------------|------------|--|------|
| | AAACCTGTCT | TACCCACTGA | GGGACGCGCT | GCGATGACGA | TTAATTGTGA | TGGTTCTAAT | 6180 |
| 5 | CCCCCTATTT | TGTAATCCAT | TAGCTTGTAA | CCCGTCTTAA | TTTGCTTCTT | AGGGCTATCG | 6240 |
| | CTGTATAACT | CTTCGACAAA | CTCCTCAACA | AACTTCTTGG | TTCCATCTTC | TTTTTTGTTA | 6300 |
| | GTAATTGTTT | TTAAATCCTT | GAGTTCATCA | ATCAAGTTGT | TAAAGTTTTG | GTTCGTAGGT | 6360 |
| 10 | TGTTGTTTGA | ACTCAGTTAC | CAATTCGTTA | GCTTTGTTGA | GCTGATAACT | TTCCAATAAT | 6420 |
| | TCTTGTTGAT | AACGTTCAAA | GAAGCCATAT | CCAATGAAAT | CGGAGTTGTA | AAGTTTAGTT | 6480 |
| 15 | ATAGTATCTG | CATCTAAAAA | TTCTTTATCT | TTAGTTGCTT | TTAAATAGAT | TTCTTGATGA | 6540 |
| | TCTATCTTTC | CGACGTCCAT | TACATAATTG | AAAAAGGTTT | TAAACTTTTC | GTTCGTAAAC | 6600 |
| | ATGTAATCTT | TAACTCTTAT | CTTTTCTAAT | ACGTCCGGTT | GTTTAAGTAG | CGTAGCGATT | 6660 |
| 20 | ATTGTACTTT | CAATTTCGAA | TTGTCCGTAA | TTCATTCGTT | TTCGCCCCCA | AATTCTGCCA | 6720 |
| | ACTTATTCAT | GAACTTATCT | AGCGCTATTT | TTCTTTGTCT | GACATATTCG | GGGTCATTCT | 6780 |
| | GCATTTTCCA | TTGGTGTGTA | GCGGTTTCGT | TATCTACTGG | CTCGATAGAT | ACTITITAG | 6840 |
| 25 | GTTCCTTACG | CATGATTGCT | GGTAAGTTAG | GCGGGTACGG | GTTGTTACTG | TTGATATAAA | 6900 |
| | CATCTACCGC | TTTTACAGTT | GGTTGATAAT | CTCCATTTTG | ACTTAATACA | TCAATCCACA | 6960 |
| | TTTCTAACTT | CGGTTTATCA | AAATCAATGT | TGTATACGTA | CCTAACTTTT | TTAATAATTT | 7020 |
| 30 | CTAATGCTTG | TGTTTTGCTC | ATCGGCATTA | GTCATCACTC | AATTCTTTTT | CCATTTGTGC | 7080 |
| | AATGACATCA | TCAGTAGTAT | TTTTTCTAGG | TGCTATTTTA | TTTTCTGCAT | CTTCTTTTGT | 7140 |
| o.c | TTTGACATTC | TCTTTAGCCC | AGTTGTTTAA | AACTTTAATT | AAATAGCCAC | CATGCGCACT | 7200 |
| 35 | TTTGCTTTTA | GTGTACTCAA | CACCTACTTT | TACAACTTCA | AAAGCGTTTG | TACCTATATC | 7260 |
| | ATCAATAGCA | AACCCTAATT | GTTCCATTTG | ATTAGGTGTT | AACTTATCAT | CCAAATTTGC | 7320 |
| 40 | AATTATATAT | TTTATTGAAG | ATGAGAAGAC | GGCTTCTCTT | TCTTCTTCTT | TATTCTTATA | 7380 |
| | TTCTTCTTCT | TTTTCTTCTT | CTCTTTCTTC | TTCTTCTTCT | GTATCGTTAC | GTAACGTTAC | 7440 |
| | GGTAACGTTA | CGTTTTGCTT | CTAGTAACTT | TTTCTGTTTC | TCACGATAGC | GTTGTTGTCG | 7500 |
| 4 5 | CAATTTATTT | TTTTCTTTAT | GCTTAGCTTT | GCTATCTAAG | CTTTGATGCT | TCTCCCAGTT | 7560 |
| | TGTCACTTTT | ATGACACCAT | TAACTTTTTC | AATCATGCCC | AATGTCTCAA | AAGTTTGAAT | 7620 |
| | TGCTAACCTT | ATTGAGTTAA | TAGGTCTATT | AAATTCATTT | GCTAACATTT | CTTCGTTGTA | 7680 |
| 50 | CGGCAAGTTT | TCGGATAGCA | TAATATAACC | TTGTTCATTG | TACTITCCTG | ATAAAGTTAG | 7740 |
| | TARCTTARCC | CARAGE | TC 1 TC CT 1 TC | morros com | **** | #1 #1 #1 #1 #1 #1 #1 #1 #1 #1 #1 #1 #1 # | |

| | CTCCTTTCAG CATTTTGTTG AGCCTCTCAT CAACTTTTAT CCACGAGTCA TGCAAGTGAT | 7920 |
|----|---|-------|
| _ | ATTTATCATC AAACGACTTA ACGCCAATTG CGTGCTGTTC ATTATGATGT TGTCTACACA | 7980 |
| 5 | GTGCTAACAC ATGTTTGTCG TAGTGATTCA TTTTGTTTCT GTTCATGCCT CTGCCGACTG | 8040 |
| | CTTCATAATG TGCCAGGTCT GCGTGAGGCT TTCCGCATAT TACACAGTTG CGGTTGATTG | 8100 |
| 10 | TAGCCCAATA TAATAACGCT TTATCTTCGC TTAACAACTT ACTCGTTTCT ACACTCATAG | 8160 |
| | GTATTTGATG ATGAAACATA AACGCTATAA TCAGTTCTAT TAACTCCCTT GCAACTTTCA | 8220 |
| | TAGAACAGTC GCGCAGACTG ATTTCTTCAT AACCTTTCAT AATTTCCAAT TCTGTTTGTA | 8280 |
| 15 | ATAATTTCT AGTTGATTCT ACTGGTTCGC CCCAGTGAAG TTCTATATCT CTACACATTG | 8340 |
| | CGAATATTTT TTTGCGTTGT TCTATAGATA GTTTTTTATT GTCCGGAACC TCTACTTCTG | 8400 |
| | CTTTTAGTGG ATATCCGTTT TCTAGTAAGT CAATGTGACT TTGTTCAAGT TCAACACCAG | 8460 |
| 20 | TAGCAACGAC GGAATAAGTA CCGTCATTGT CTTTCTGGTA TCTTGTAATG TATTGCATTT | 8520 |
| | AAACCACGTC CTAGAACGGT AAATCATCAT CATTGATTTC TATTGGACCA TTAGCATTAG | 8580 |
| | CGAATGGGTT TGATTGTTGA CTCATTGGCG TCTGTTTCCC ATTTGCTTGC TGTTCTTTTT | 8640. |
| 25 | GTTTCATCTC ATCAGTTTTA GGTTCTGGTT TATTAACTAC TTCATCGTCT TTATTCCAAA | 8700 |
| | CTTTTACATA TGAGAGTCTT ACAAAATACT TGCCTTGTTC CTCGTTAAAT TTATTTTTAA | 8760 |
| | GTACAATAGT TCCGATTTTG TTAATTAATT GATCTGTGTC AAAAGTTAAA TCTGGTAAGT | 8820 |
| 30 | TCAATTTAAT TCCTAATCTA CTAAGTAACT CGATATATTG TTTTTCTTGA TAATCTTGTT | 8880 |
| | GGAATGGTGG GACGAATTGG TTGTGTTTGT ATTGTTTACC TTCGTTGTTT TCAAAAACAA | 8940 |
| 35 | TCGTGAAGTA TCTGTTTTCT CTGTCGTTAA ACTCGACATT TGCAACTTTT ACTGTAAATT | 9000 |
| | CTCCAGCTCC TAAAAAGTCC CCACCTTTCA TGAATGCCTC TTGATTAGTT TCTTGAATGT | 9060 |
| | ATTGTGTTCT ACCAGTGATT TTCATAATTT TTATACCGTC CTTTTAATTA ATTTTTAATT | 9120 |
| ю | ACCATTTCTA ATTGCTTGTA CAACATCGTT AATACTTGGA TTAATGAAAC GTTTGTTGTT | 9180 |
| | AATTTTGATG TTGCTTGAGT GTCTTATCTT TGTCTCGAAT AAATTTGATG GTTCAGCGTT | 9240 |
| | AAGTACATAT TGATAAGTTT TTTCGCCGTC TTGCTCATGT TCTTCTATTG TCATTCTTGC | 9300 |
| 5 | TAACACGTCA GATTGACTGA TGACTGCTTT TTTTATTTGG TCTTGTGCCT CTATCGTGAT | 9360 |
| | TGTTGGATTG ATAGTACTTC CCTCATCATC TTTGTCTTTG TTAATGCCCT CGTGTCCGCT | 9420 |
| | TATAGCAAGA TGAAATTGAT AATGTTCTTG TAATTTAGAA ATATAACGAT AAATACTTAC | 9480 |
| o | AATGCGTGTA GCACACTCGC CCCAATCATT AAATGTCGGT TTCTTTGATT TACCGTCCAT | 9540 |
| | GATGTCGTCC ATAGTGATAT CACGTAACTT TTGGATTGTT TCAATCACTA CAACATCAAT | 9600 |

AAAATGCTTA TAATTCTTAA TCTGCACAAC TGCCCCATCT TCTGTTACCG TTGTTCCGTC 9720 CTCATTTATA TCTAGTACTA AGGCATTGTT ATCTTTTGTT AAAAACGTAG TTTTACCAGT 9780 ACCGAACTTG CCGTATATCG CAAATTTATA AAACTTGTTT GCATTTTGTT TGCTGATGTC 9840 TITTACACCT AGTTGCGTTA AAATATCGAC ATCTTGATTA GTTTTTTCAG TCATCTATTC 9900 TCCCACCTTT ACCGTGTATG ACGTTGGTTT CTCCACAATG CTAGCACCCT CTAAAACTTC 9960 GCCGTTTGCG TCAATCAATG TGCCGTTTTC AGTTACATTG AAATCTTTCT TAATGTCTGA 10020 TTGGCTAAGT TTTTTAGTTA CTTTTACATA GTTGTCAAAA CCTCGTTGCT CAAGTTGTnT 10080 AATGACTTCT TGCTCATTGC TAACTTGAAT GACTTTTGAA CCTTTTCTGG CTGTCACTTT 10140 TCCGTAAGLG TATTCAACTT GAATTTGCTA TCTTGTTCTT TTTGTATTCT GTAATATTCA 10200 ATTACAAGGC TTTGTAAATA TTCTTTGCCA CTCTGTAATT TTTCTACTTC TTTATCTTTC 10260 CATTCGTTTA TGCGTTCAAT TTCTTTATTT GCTAAATCGT TGATTTCATT CTCTTTAGTT 10320 GTGATTGCAT CCAGTTTCTn AAAAACCCAG TTAGCACTGT CTAGATCAGT nACTTTGAAT 10380 CGGTCGTCTT GTTCGAATGT n 10401

25 (2) INFORMATION FOR SEQ ID NO: 150:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2989 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 150:

TTTCTCTCTA TTATTCTCGA TGCGTAGATA ATTGTTTAAA TTTAAGTTTA TAGTAATGTT 60 GAGTITATAA TITCATATAT CTAAAAACAG GTGTTGTATA TATAATCATT CATCTAGTTA 120 TACTTACTTT AAAAATAATA TAATTTCATG CGATGCAATT CATTGATGGA TGTTTTTAAT-180 CTTAATCAAA TCCAAATAAA GCATATATTT TTAAATTCAC TTTCTTTCGA ATCGATTTTT 240 ATCTCTTGNA TTAAACTTTT CCATTGTTTC ATTAAAGCTC TCTGTCATAT CTATTCCCAT 300 TGAATTCGCT AAACATAACA ACACAAATAA ATTATCACCT AATTCTGCTT TAATCGTATT 360 TGCTTCCTCT GAATCTTTCT TCTTTTTTTC ACCATAGGTA TGATTTATTT CACGTGCAAG 420 TTCGCCCACT TCTTCAGTCA ATCTAGCTAA GTTAGCTAAT GGTGAAAAAT ATCCTGTTTT 480 AAATTGTCCA ATATATTCAT CAACTTCACG TTGCATTTCT ACCATTGATT TCATTTCTAC 540 GTTCTCCTTA TATTGCATTT CTAATATAGT ATATATCAAT TTGAAGTCTC ATGCATGTTT 600

55

50

5

10

15

20

30

35

40

| | AATTCAGTTT ATATAAAT | GT AATGCATTC | C TAACTAAATT | AAATCAATTO | AAATTGGGAT | 72 |
|----|----------------------|---------------|--------------|------------|------------|------|
| _ | TATAACTTTA TGATACGT | AC CACTACAAT | A AAATAATATA | GTGAATAATC | TACCATTAGA | 78 |
| 5 | AAAATAAGCA CAAAAAA | CT AGCAACCAC | A CAAAAATGTG | ATTAGCTAGT | TAATAAGTGT | 84 |
| | CTAATTTAAG TTAATTGT | TA ATCTATAAGI | A TTAATCACTT | GAACGCGCAA | TCAAAATAAT | 900 |
| 10 | ACGTACAAGC TCTGCTAC | AG CGACTGCAG | T TGCTGCAACA | TAAGTCATTG | CTGCTGCAGA | 960 |
| ,, | TAATACTTTA CGCGCATG | CT TGTATTCTT | TTCATTTACA | ATGTTCAATG | CCGTAATTTG | 1020 |
| | TTTCATCGCT CTTGAACT | CG CATCAAACTO | AACTGGTAAC | GTAACAATTG | AGAATAATAC | 1086 |
| 15 | CGCTAATGAC ATTAAACC | AG CACCAATCC | TAAAGCAGTT | GAACCAAATG | CACTACCTAT | 1140 |
| | CGCTGTTAAG ATAATACC | IA ACATGATGAT | CATATAACTT | AATGAACTCC | CTAGGTTTGC | 1200 |
| ٠ | AACAGGTACT AATGCTGC | IC TGAATCTTAA | GAACCAATAT | CCTTGGTGAT | CTTGAATGGC | 1260 |
| 20 | ATGACCAACT TCGTGGGC | IG CAATTGCAGI | TCCAGCAACT | GATGGTCTGT | CATAGTTTGC | 1320 |
| * | AGGAGATAGT GAAACAAC | TT TCTTTTTAGG | ATCGTAATGA | TCTGTTAAGA | ATCCTTCACC | 1380 |
| | TITAACAACT TCGACATC | AT AAATACCGTT | TGCATGTAAA | ATTTCTAATG | CAACTTCACG | 1440 |
| 25 | ACCCGTTTTA CCACTAGT | rg atctaacttg | TGAATATTTC | TCATAGTTAG | ATTTAACTTT | 1500 |
| • | GTGTTGTGCC CATAAAGG | NA GCACCATTAA | TATTACGAAA | TAAATTATCA | TAGTAAAAAT | 1560 |
| | TGAAGACAAT AAACTCAC | C TCCTTTATAA | ATATTTTACT | GTCATTTGCC | GTTTTTATCA | 1620 |
| 30 | AATCATTTAC ACTTTAATA | A TTTGTTTAAT | TCAATATAAA | GCAAAAGTCC | AAAAACACTT | 1680 |
| | AGACAACATG ATAATACAG | C AATTTGCCAC | ACATGTGTAG | TTATAAAATC | ATAATATGGA | 1740 |
| | AATTGAAGGT GAAAATAGT | C AATATAATCA | TTCAAAAACA | CCCAAATCAT | yGCTACACTG | 1800 |
| 35 | ATTCCAATCA TAGAACGTT | T AAACCTAGGA | TAGAAGTAAA | TTGCCTGAAC | AGCCATTATA | 1860 |
| - | CTGTGGGAAA ACATTAATA | C CAAACCATTT | ACTGTAATAT | CACCTTGTTC | TAAATAATAA | 1920 |
| 10 | AATATATTCA TTATAACTC | C CCAAATCCCA | TATTTGAATA | ATGTTACAAA | TGCCAGTGCA | 1980 |
| | TCGATAATAC TATTTTGTT | T TTGAATTAAT | ATCAATGAGA | TAGAAATAAC | TAAGTATAAT | 2040 |
| | ATTGCAGTTG GGCTATCTG | G AACAAAAATC | TTAAAATGCC | AGGGCGTATG | ACTTAATTGT | 2100 |
| 5 | TCACCATACC ATATATAAC | C ATAAATCATC | CCTAATATAT | TACAAATGAG | TAGCATCATT | 2160 |
| | AACCAAGAAC GTTGATAAA | G TGTATATTGC | CAAAATGCTT | TAATTGTCAT | CTGCTAAGTC | 2220 |
| | CTCAAATTGA TTATGTTTA | T TTACTAGCTT | GAGTGTATTT | AAAATTTGCG | TTAGTTGATA | 2280 |
| o | AAAACGTTGC TTTTCATTC | A TCTGTAAACT | TAAATCAATA | TTGTGTAACA | AGTAATCTAT | 2340 |
| | TAATAACGCA TGTTTATGC | C GATCTATAGC | CATACTATTT | AAGTCATGAA | GATAAGTTTG | 2400 |

| TGACACGTTT | GCGAAGTGAA | TTTGAATATC | AAAAGCACAG | TTATGATTAG | CGATATAATC | 2520 | | | | |
|------------|------------------------------------|------------|------------|------------|------------|------|--|--|--|--|
| AAATATTTCA | TTTGTATTCA | TTAACTTTAT | ATTACGCTTA | GTAAATTGAA | TTGCAGAAGC | 2580 | | | | |
| GTGACTTCCC | ACTTCTGCAA | TTTCTAATGT | TTCATGATGA | TTAATTTTTG | TATCTACAAA | 2640 | | | | |
| ATGAATGTTT | GCCAATTTCG | CCTCATTCAC | TTTTATATAG | TTAAGCACCC | AAACTGCAAT | 2700 | | | | |
| ACGCGACTTA | AATCGATATT | GAAAAAGTAA | ATATTCAATA | AAACTTTCTT | TAATTTGATT | 2760 | | | | |
| GAGTGTCTCT | GACATCAAAT | ACCCCATTTT | AAGATTGCAA | TCTTGaTAAT | TCGTCATGCC | 2820 | | | | |
| AATTTTCGTT | ACTTGGCTCT | AGTTCCAACA | ATTGATTTAA | AATAGTAATT | GCTTGTTCCT | 2880 | | | | |
| TTTGACCAAT | TTCAATTAAA | TAGAAATAAT | AATCACTCAT | AAAATCAATA | TTTGTTTTCA | 2940 | | | | |
| TCGTTGGATA | TGCTAATTCA | AAGAAATGTT | GAGCTTCTTT | ATCTCGCTC | | 2989 | | | | |
| (2) INFORM | 2) INFORMATION FOR SEQ ID NO: 151: | | | | | | | | | |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1143 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 151:

| CATCAAC. | ıcc | TIAATIACAC | IGIAAAIGAI | AIGCGICIII | IIGACAACIA | TATTTGTCAA | 60 |
|----------|-------------|------------|------------|------------|------------|------------|-----|
| ATCTACAG | CA | AAAAATATGA | TTATCCACCT | ATGTATGACA | TTTTGAAACA | AACACCTCAA | 120 |
| CGCCTAC | AA G | TCATAATTGT | TTACTTTCGT | TACACCTTCC | TGCATAATTA | ACAGCATTCT | 180 |
| AATTTTAA | STA | TGATGCACGC | ATTTTCACTA | AATCAAACCA | TTCAAAGGAG | ACTATTATGG | 240 |
| CATTTAC | ATT | ATCTGCAATT | CAACAAGCAC | ATCAACAATT | TACTGGTGTT | GACTTTCCAA | 300 |
| AACŢĀTT | CAA | AGCTTTTAAA | GATATGGGGA | TGACTTACAA | TATCGTCAAC | ATTCAAGATG | 360 |
| GCACTGC | AAC | ATACGTACAT | CAATCAGAAG | ATGATATCGT | TACGTCATCT | GTAAAAAGTA | 420 |
| ATCATCC | TGT | TGCTCAAAAA | TCAAACAAAA | CAATAGTTCA | AGACGTCTTA | ACTAGACATC | 480 |
| AACAAGG | GCA | AACAGATTTT | GAAACATTTT | GTGATGAAAT | GGCTGAAGCT | GGCATTTATA | 540 |
| AATGGCA | TAT | CGATATTCMA | GCGGGCACTT | GTACTTATAT | CGACTTGCAA | GACCAAGCTG | 600 |
| TTATTTC | AGA | ATTAATCCCT | CAATAAACTA | TATTTATAGC | AACATTTTAA | TTATTTCATA | 660 |
| AAATTTT | ATT | GATAATCATT | ATCGTTCGGT | ATAAAGTAAA | TACTATATAC | TACTTATGAG | 720 |
| TGAGGTT | GAT | TATCATGATA | ACTAACACTT | TTATTTTAGG | CATCACAGGC | CCAACAAGTC | 780 |
| TTCTCT | тат | TAGCATTATC | ССТТААТТА | THEFT | CAAAAAATTA | CCACAATTTG | 840 |

AGTCTCACGA TACACCCAGT AAGGAATCGA AACAACAGCG AGAGCAATAG CACTGACCAC 960
ACCTTACTGG TTCACTTTAG CGAACTACGC CATCGGTTAG TAAAAATTTT ATTGTCGTTC 1020
GTCATTACGG TCATCGTCGT ATATGTYTCA TCATTTTGGT GGATGACACC ATTCATAACG 1080
TATATYACCC GGCACATGTG TCCTTACATG CATTTCATTC ACAGAAATGA TACAAATAAC 1140
GTG

(2) INFORMATION FOR SEQ ID NO: 152:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7953 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 152:

| CAACGCCTGA A | ACGTAAACCA | TATCGTTTCG | CGATTTCCTC | ATCTTGACTA | TTTACTAAAA | 60 |
|---------------|-------------------|------------|------------|------------|------------|------|
| ACTCTCTCAT G | GCGATTAAT | GTTTCTTTTT | CTTCTTTAGT | TAATGGTAAT | TCTAACTCAG | 120 |
| CTGCTTTTTG A | ACGCAAAGTT | GGATGACCAT | CTCTAATGAT | GTCTTTCATT | GTTAACATAT | 180 |
| ATTGCACCTT C | CCTTATTTTA | ATTTGTTTTA | GTTGAATGAC | AGTAAAAAGG | TTGTTAAGAT | 240 |
| ACTCATACAT I | TTTATGTGT | AAATATCTAC | AAAGTTAACC | AACTACTGCC | AATGTTTATT | 300 |
| TTAGATAGTA I | PATGTAAATT | TTCAaGAtAT | GCgTAATTGC | gttaaaaaat | Gattaaagtg | 360 |
| TTGGTTTCAA G | GCAATGATAC | TTTAGAAATT | TATTTATCAT | CTTGACTTTA | AAAATTATAT | 420 |
| TATAAATGAC G | STAACTGTCA | ACAGATATAC | TTAGTArTGA | AGATGTGTAA | TGTAATTGTT | 480 |
| TAAAATTGAT 1 | TTCCAAGCAG | ATTTTATTTA | TCATTTAATT | TAAATAGCAA | GTGGAGGTAC | 540 |
| AAGŢĀATGAA A | TTTGGAAAA | ACAATCGCAG | TAGTATTAGC | ATCTAGTGTC | TTGCTTGCAG | 600 |
| GATGTACTAC G | GATAAAAA | GAAATTAAGG | CATATTTAAA | GCAAGTGGAT | AAAATTAAAG | 660 |
| ATGATGAAGA A | CCAATTAAA | ACTGTTGGTA | AGAAAATTGC | TGAATTAGAT | GAGAAAAAGA | 720 |
| AAAAATTÄÄC, I | TGAAGATGTC | aatagtaaag | ATACAGCAGT | TCGCGGTAAA | GCAGTAAAGG | 780 |
| ATTTAATTAA A | VAATGCCGAT | GATCGTCTAA | AGGAATTTGA | AAAAGAAGAA | GACGCAATTA | 840 |
| AGAAGTCTGA A | CAAGACTTT | AAGAAAGCAA | AAAGTCACGT | TGATAACATT | GATAATGATG | 900 |
| TTAAACGTAA A | AGAAGTAAÁA | CAATTAGATG | ATGTATTAAA | AGAAAAATAT | AAGTTACACA | 960 |
| GTGATTACGC G | GAAAGCATaT | AAAAAGGCTG | TAAACTCAGA | GAAAACATTA | TTTAAATATT | 1020 |
| TAAATCAAAA 1 | rgacgcgaca | CAACAAGGTG | TTAACGAAAA | ATCAWAAGCA | ATAGAACAGA | 1080 |

55

5

10

15

20

25

30

35

40

45

| | AAGAAAAGCA | AGACGTTGAT | CAATTTAAAT | AATTAATATA | ATACAGATGG | TAGGAAACAA | 1200 |
|----|------------|------------|------------|-------------|------------|--------------|------|
| | CTAATACAGT | TCCTATTATC | TGTATCTTTT | TTTATTAAAA | CAGAACTTTT | TCAAATGGTT | 1260 |
| 5 | TAACAGTCCC | ATTTATTTGT | GGTACAATTA | GTAAGGATAA | AATGAATTTC | TATACAATTA | 1320 |
| | TGGGAAAGGT | ATTGTGAATT | GAATGGCTCC | TAAGTTACAA | GCCCAATTCG | ATGCAGTAAA | 1380 |
| • | AGTTTTAAAT | GATACTCAAT | CGAAATTTGA | AATGGTTCAA | ATTTTGGATG | AGAATGGTAA | 1440 |
| 0 | CGTCGTAAAT | GAAGACTTAG | TACCTGATCT | TACGGATGAA | CAATTAGTGG | AATTAATGGA | 1500 |
| | AAGAATGGTA | TGGACTCGTA | TCCTTGATCA | ACGTTCTATC | TCATTAAACA | GACAAGGACG | 1560 |
| 5 | TTTAGGTTTC | TATGCACCAA | CTGCTGGTCA | AGAAGCATCA | CAATTAGCGT | CACAATACGC | 1620 |
| | TTTAGAAAAA | GAAGATTACA | TTTTACCGGG | ATACAGAGAT | GTTCCTCAAA | TTATTTGGCA | 1680 |
| | TGGTTTACCA | TTAACTGAAG | CTTTCTTATT | CTCAAGAGGT | CACTTCAAAG | GAAATCAATT | 1740 |
| 0 | CCCTGAAGGC | GTTAATGCAT | TAAGCCCACA | AATTATTATC | GGTGCACAAT | ACATTCAAGC | 1800 |
| | TGCTGGTGTT | GCATTTGCAC | TTAAAAAACG | TGGTAAAAAAT | GCAGTTGCAA | TCACTTACAC | 1860 |
| | TGGTGACGGT | GGTTCTTCAC | AAGGTGATTT | CTACGAaGGT | ATTAACTTTG | CAGCAGCTTA | 1920 |
| 25 | TAAAGCACCT | GCAATTTTCG | TTATTCAAAA | CAATAACTAT | GCAATTTCAÄ | CACCAAGAAG | 1980 |
| | CAAGCAAACT | GCTGCTGAAA | CATTAGCTCA | AAAAGCAATT | GCTGTAGGTA | TTCCTGGTAT | 2040 |
| | CCAAGTTGAT | GGTATGGATG | CGTTAgcTGT | nATATCAAGC | AACTAAAGAA | GCACGTGACC | 2100 |
| 30 | GCGCAgTTGC | AGGTGAAGGT | CCAACATTAA | TTGAAACTAT | GACATATCGT | TATGGTCCTC | 2160 |
| | ATACAATGGC | TGGTGACGAT | CCAACTCGTT | ACAGAACTTC | AGACGAAGAT | GCTGAATGGG | 2220 |
| 35 | AGAAAAAAGA | CCCATTAGTA | CGTTTCCGTA | AATTCCTTGA | AAACAAAGGT | TTATGGAATG | 2280 |
| .5 | AAGACAAAGA | AAATGAAGTT | ATTGAACGTG | CAAAAGCTGA | TATTAAAGCA | GCAATTAAAG | 2340 |
| | AGGÇTGATAA | CACTGAAAAA | CAAACTGTTA | CTTCTCTAAT | GGAAATTATG | TATGAAGATA | 2400 |
| 10 | TGCCTCAAAA | CTTAGCAGAA | CAATATGAAA | TTTACAAAGA | GAAGGAGTCG | AAGTAAGCCA: | 2460 |
| | TGGCACAAAT | GACAATGGTT | CAAGCGATTA | ATGATGCGCT | TAAAACTGAA | . CTTAAAAATG | 2520 |
| | ACCAAGATGT | TTTAATTTTT | GGTGAAGACG | TTGGTGTTAA | CGGCGGTGTT | TTCCGTGTTA | 2580 |
| 15 | CTGAAGGACT | ACAAAAAGAA | TTTGGTGAAG | ATAGAGTATT | CGATACACCT | TTAGCTGAAT | 2640 |
| | CAGGTATTGG | TGGTTTAGCG | ATGGGTCTTG | CAGTTGAAGG | ATTCCGTCCG | GTTATGGAAG | 2700 |
| | TACAATTCTT | AGGTTTCGTA | TTCGAAGTAT | TTGATGCGAT | TGCTGGACAA | ATTGCACGTA | 2760 |
| 50 | CTCGTTTCCG | TTCAGGCGGT | ACTAAAACTG | CACCTGTAAC | AATTCGTAGC | CCATTTGGTG | 2820 |
| | GTGGCGTACA | CACACCAGAA | TTACACGCAG | атаасттаса | AGGTATTTA | GCTCAATCTC | 2886 |

| | CTATTAGAAG | TAATGACCCA | GTCGTATACT | TAGAGCATAT | GAAATTGTAT | CGTTCATTCC | 3000 |
|------------|------------|------------|------------|------------|------------|-----------------|------|
| | GTGAAGAAGT | ACCTGAAGAA | GAATATACAA | TTGACATTGG | TAAGGCTAAT | GTGAAAAAG | 3060 |
| 5 | AAGGTAATGA | CATTTCAATC | ATCACATACG | GTGCAATGGT | TCAAGAATCA | ATGAAAGCTG | 3120 |
| | CAGAAGAACT | TGAAAAAGAT | GGTTATTCTG | TTGAAGTAAT | TGACTTACGT | ACTGTTCAAC | 3180 |
| 10 | CAATCGATGT | TGACACAATT | GTAGCTTCAG | TTGAAAAAAC | TGGTCGTGCA | GTTGTAGTTC | 3240 |
| 10 | AAGAAGCACA | ACGTCAAGCT | GGTGTTGGTG | CAGCAGTTGT | AGCTGAATTA | AGTGAACGTG | 3300 |
| | CAATCCTTTC | ATTAGAAGCA | CCTATTGGAA | GAGTTGCAGC | AGCAGATACA | ATTTATCCAT | 3360 |
| 15 | TCACTCAAGC | TGAAAATGTT | TGGTTACCAA | ACAAAAATGA | CATCATCGAA | AAAGCAAAAG | 3420 |
| | AAACTTTAGA | ATTTTAATAC | ATTTTAAAAG | TTAACGAAGT | TAGCGTATTT | TAGTCTCATT | 3480 |
| | GATTAAAATG | AAATGTTTAA | TTTACGAAAT | CTTAGGAGGG | CAAAAACGTG | GCATTTGAAT | 3540 |
| 20 | TTAGATTACC | CGATATCGGG | GAAGGTATCC | ACGAAGGTGA | AATTGTAAAA | TGGTTTGTTA | 3600 |
| | AAGCTGGAGA | TACTATTGAA | GAAGACGATG | TTTTAGCTGA | GGTACAAAAC | GATAAATCAG | 3660 |
| | TAGTAGAAAT | CCCATCACCA | GCATCTGGTA | CTGTAGAAGA | AGTTATGGTA | GAAGAAGGTA | 3720 |
| 25 | CAGTAGCTGT | AGTTGGTGAC | GTTÄTTGTTA | AAATCGATGC | ACCTGATGCA | GAAGATATGC | 3780 |
| | AATTTAAAGG | TCATGATGAT | GATTCATCAT | CTAAAGAAGA | ACCTGCGAAA | GAGGAAGCGC | 3840 |
| | CAgcAGaGCA | AGCACCTGTA | GCTACTCAAA | CTGAAGAAGT | AGATGAAAAC | AGAACTGTTA | 3900 |
| 3 <i>0</i> | AAGCAATGCC | TTCAGTACGT | AAATACGCAC | GTGAAAAAGG | TGTTAACATT | AAAGCAGTTT | 3960 |
| | CTGGATCTGG | TAAAAATGGT | CGTATTACAA | AAGAAGATGT | AGATGCATAC | TTAAATGGTG | 4020 |
| 35 | GTGCACCAAC | AGCTTCAAAT | GAATCAGCTG | CTTCAGCTAC | AAGTGAAGAA | GTTGCTGAAA | 4080 |
| ,5 | CTCCTGCAGC | ACCTGCAGCA | GTAACATTAG | AAGGCGACTT | CCCAGAAACA | ACTGAAAAA | 4140 |
| | TCCCTGCTAT | GCGTAGAGCA | ATTGCGAAAG | CAATGGTTAA | CTCTAAGCAT | ACTGCACCTC | 4200 |
| 10 | ATGTAACATT | AATGGATGAA | ATTGATGTTC | AAGCATTATG | GGATCACCGT | AAGAAATTTA | 4260 |
| | AAGAAATCGC | AGCTGAACAA | GGTACTAAGT | TAACATTCTT | ACCITATGIT | GTTAAAGCAC | 4320 |
| | TIGITICIGC | ATTGAAAAAA | TACCCAGCAC | TTAACACTTC | ATTCAATGAA | GAAGCTGGTG | 4380 |
| 15 | AAATCGTTCA | TAAACATTAC | TGGAATATCG | GTATTGCAGC | AGACACTGAT | AGAGGATTAT | 4440 |
| • | TAGTACCTGT | TGTTAAACAT | GCTGATCGTA | AGTCTATTTT | CCAAATTTCA | GATGAAATTA | 4500 |
| | ATGAATTAGC | TGTTAAAGCA | CGTGATGGTA | AATTAACAGC | CGATGAAATG | AAAGGTGCTA | 4560 |
| 50 | CATGCACAAT | CAGTAATATC | GGTTCAGCTG | GTGGACAATG | GTTCACTCCA | GTTATCAATC | 4620 |
| | ACCCAGAAGT | AGCAATCTTA | GGAATTGGCC | GTATTGCTCA | AAAACCTATC | CTTS S S C S TC | 4600 |

| | ATGGTGCAAC | TGGCCAAAAT | GCAATGAATC | ACATTAAACG | TTTATTAAAT | AATCCAGAAT | 4800 |
|----|------------------------|--------------|------------|-------------|--------------|------------|------|
| | TATTATTAAT | GGAGGGGTAA | AACATGGTAG | TTGGAGATTT | CCCAATTGAA | ACAGATACTA | 4860 |
| 5 | TAGTAATCGG | AGCAGGTCCT | GGTGGATACG | TTGCAGCAAT | TCGTGCAGCT | CAATTAGGAC | 4920 |
| | AAAAAGTAAC | AATCGTTGAG | AAAGGTAATC | TTGGTGGTGT | TTGCTTAAAC | GTAGGATGTA | 4980 |
| 10 | TTCCTTCAAA | AGCATTACTA | CATGCTTCTC | ACCGTTTTGT | TGAAGCACAA | CATTCTGAAA | 5040 |
| | ACTTAGGTGT | TATTGCTGAA | AGTGTTTCTT | TAAACTTCCA | AAAAGTTCAA | GAATTCAAAT | 5100 |
| | CATCAGTTGT | TAATAAATTA | ACTGGTGGTG | TTGAAAGCTT | ACTTAAAGGT | AACAAAGTTA | 5160 |
| 15 | ACATCGTTAA | AGGTGAAGCA | TATTTCGTAG | ATAACAATAG | CTTACGTGTT | ATGGACGAAA | 5220 |
| | AGAGCGCACA | AACATACAAC | TTTAAAAATG | CAATCATTGC | AACAGGTTCA | AGACCAATTG | 5280 |
| | AAATTCCTAA | TTTCAAATTC | GGTAAACGTG | TTATCGACTC | AACAGGTGCT | TTAAACTTAC | 5340 |
| 20 | AAGAAGTACC | aGGTAAATTA | GTTGTAGTTG | GTGGAGGATA | CATTGGATCA | GAATTAGGTA | 5400 |
| | CAGCATTTGC | TAACTTTGGT | TCAGAAGTAA | CCATCCTTGA | AGGTGCTAAA | GATATCTTAG | 5460 |
| | GTGGCTTCGA | AAAACAAATG | ACACAACCTG | TTAAAAAAGG | TATGAAAGAA | AAAGGTGTTG | 5520 |
| ?5 | AAATCGTTAC | TGAAGCTATG | GCTAAATCAG | CTGAAGAAAC | AGATAACGGA | GTTAAAGTTA | 5580 |
| | CTTATGAAGC | TAAAGGCGAA | GAGAAAACAA | TCGAAGCTGA | TTATGTATTA | GTAACTGTAG | 5640 |
| | GTCGTCGTCC | AAACACAGAC | GAATTAGGCC | TAGAAGAATT | AGGTGTTAAA | TTCGCTGACC | 5700 |
| 30 | GTGGATTATT | AGAAGTTGAT | AAACAAAGCC | GTACGTCTAT | CAGCAATATC | TATGCAATTG | 5760 |
| | GTGATATCGT | TCCAGGTTTA | CCACTTGCTC | ACAAAGCTAG | CTATGAAGCT | AAAGTTGCTG | 5820 |
| 35 | CTGAAGCAAT | TGATGGTCAA | GCTGCTGAAG | TTGATTACAT | TGGTATGCCA | GCAGTATGCT | 5880 |
| | TTACTGAACC | AGAATTAGCT | ACAGTTGGTT | ATTCAGAAGC | GCAAGCTAAA | GAAGAAGGTT | 5940 |
| | TAGÇAATTAA | AGCTTCTAAA | TTCCCATATG | CAGCAAATGG | TCGTGCATTA | TCATTAGATG | 6000 |
| 40 | ATACTAACGG | ATTTGTTAAA | CTTATTACAC | TTAAAGAAGA | TGATACTTTA | ATCGGTGCTC | 6060 |
| | AAGTAGTTGG | TACTGGTGCA | TCAGATATTA | TCTCTGAATT | AGGTTTAGCA | ATTGAAGCTG | 6120 |
| | GTATGAATGC | TGAAGATATC | GCATTAACAA | TCCATGCACA | TCCAACATTA | GGTGAGATGA | 6180 |
| 45 | CTATGGAAGC | AGCAGAAAAA | GCTATCGGAT | ACCCAATCCA | TACAATGTAA | TAACTGATTA | 6240 |
| | TCTATAAAGA | TTCAGTCATT | AAAAGCTGTA | GCATATGCTA | CGGCTTTTTT | GTTTTAGGTA | 6300 |
| | AAGTAATGTA | AGGAAATTGA | TTTGAGATAT | CGTTAACATG | TGACATGCAT | GTTATACTAG | 6360 |
| 50 | CGATGCTAAT | AAAAGAATTG | AAATGGAGGG | TTCAACAATG | GAATATGAGT | ATCCAATTGA | 6420 |
| | ********************** | 30T3 3TC3 3C | NCATCATTTC | ACTCATA AAT | חדע מייידיים | ATCTACACAA | 6480 |

| | MATIGIGCCT | GCTAAAGCAG | AGGAAAAAC | AATTTTTAA1 | ACTITICGAA | A AAAGTAGTGG | 6600 |
|----|--------------|------------|------------|--------------|--------------|--------------|------|
| | CTATAATAGT | TACAAAGCAG | TTCAAGATGT | T AAAAACTCAC | TCTGAAGAAC | AAAGAGTAAC | 6660 |
| 5 | AGCTAAAnaa | TAATTCGTTC | GAAATTAACA | CAATTTAATA | GGAATTTTTC | TTTAAAACTA | 6720 |
| | TTGCTAATAA | AGCTATATTT | TGATACCTTT | ATCAAGTGTT | ' AAACAAAATG | TTTGATAAAA | 6780 |
| 10 | GTAAACTTAA | TATAGCTTTT | TTAGGTGGAA | AAATAAATGA | ACATAGGTAA | TAAAATTAAA | 6840 |
| | | | | | | AGACTTATCG | 6900 |
| | | | | | | AACTTTCTTA | 6960 |
| 15 | | | | | | TGAAAATGAA | 702Ó |
| | • | | | • | | AGGTTATATA | 7080 |
| • | | | | | | ATTAACTTTA | 7140 |
| 20 | AAGCCTGGAG | | • | | | • | 7200 |
| | ATGTCAGGTC . | | | ~ | | | 7260 |
| | GTTTTGTATT | | | | | • | 7320 |
| ?5 | CGAATACTTA ' | | | | | | 7380 |
| | TTATCATTAA I | | | | | | 7440 |
| 30 | ATTGATATTG I | | | | | | 7500 |
| | ACAATTTTAA 1 | | | | | | 7560 |
| | AACAAACCAA | | | | | | 7620 |
| 5 | TATGCATTAT 1 | | | | | | 7680 |
| | AAATTATCAA A | | | | | | 7740 |
| | TCAGGTTATG A | | | | | | 7800 |
| 0 | ATTGCACGTG C | | | | | | 7860 |
| | TTAGATTTGA A | | | - | GAGAATTGCa | ATCTAGATTA | 7920 |
| | GGLATTACAT T | TATATTTGT | AACACATGAT | CCA | • . | | 7953 |

(2) INFORMATION FOR SEQ ID NO: 153:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2347 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

50

| | GGCGTGATCA | TACGACCGTC | ATTCATGCTC | ATGAAAAAAT | ATCTAAAGAT | TTAAAAGAAG | 60 |
|----|------------|------------|------------|------------|------------|------------|------|
| | ATCCTATTTT | TAAACAAGAA | GTAGAGAATC | TTGAAAAAGA | AATAAGAAAT | GTATAAGTAG | 120 |
| 5 | GAAACTTTGG | GAAATGTAAT | CTGTTATATA | ACAGCACTAA | TGATNACAAT | CATTTTTTAC | 180 |
| | ATTTCTATAT | GCTAATGTGG | CAAGATGAGC | AAAACTCATT | TTGTGGATaA | TGTTTaAAAG | 240 |
| | TCATACACAC | CATACACAAG | TTATCAACAT | GTGTATAAYT | cGcCAAATCT | ATGTTTTTAA | 300 |
| 10 | GACTTATCCA | CCAATCCACA | GCACCTACTA | СТАТТАСТАА | GAACTTAAAA | CCTATATAAT | 360 |
| | TATATATAAA | CGACTGGAAG | GAGTTTTAAT | TAATGATGGA | ATTCACTATT | AAAAGAGATT | 420 |
| 15 | ATTTTATTAC | ACAATTaAAT | GACACATTAA | AAGCTATTTC | ACCAAGAACA | ACATTACCTA | 480 |
| | TATTAACTGG | TATCAAAATC | GATGCGAAAG | AACATGAAGT | TATATTAACT | GGTTCAGACT | 540 |
| | CTGAAATTTC | AATAGAAATC | ACTATTCCTA | AAACTGTAGA | TGGCGAAGAT | ATTGTCAATA | 600 |
| 20 | TTTCAGAAAC | AGGCTCAGTA | GTACTTCCTG | GACGATTCTT | TGTTGATATT | TAAAAAATA | 660 |
| | TACCTGGTAA | AGATGTTAAA | TTATCTACAA | ATGAACAATT | CCAGACATTA | ATTACATCAG | 720 |
| | GTCATTCTGA | ATTTAATTTA | AGTGGCTTAG | ATCCAGATCA | ATATCCTTTA | TTACCTCAAG | 780 |
| 25 | TTTCTAGAGA | TGACGCAATT | CAATTGTCGG | TAAAAGTGCT | TAAAAACGTG | ATTGCACAAA | 840 |
| | CAAATTTTGC | AGTGTCCAcC | TCAGAAACAC | GCCCAGTACT | AACTGGTGTG | AACTGGCTTA | 900 |
| | TACAAGAAAA | TGAATTAATA | TGCACAGCGA | CTGACTCACA | CCGCTTGGCT | GTAAGAAAGT | 960 |
| 30 | TGCAGTTAGA | AGATGTTTCT | GAAAACAAAA | ATGTCATCAT | TCCAGGTAAG | GCTTTAGCTG | 1020 |
| | AATTAAATAA | AATTATGTCT | GACAATGAAG | AAGACATTGA | TATCTTCTTT | GCTTCAAACC | 1080 |
| 35 | AAGTTTTATT | TAAAGTTGGA | AATGTGAACT | TTATTTCTCG | ATTATTAGAA | GGACATTATC | 1140 |
| 55 | CTGATACAAC | ACGTTTATTC | CCTGAAAACT | ATGAAATTAA | ATTAAGTATA | GACAATGGGG | 1200 |
| | AGTITTATCA | TGCGATTGAT | CGTGCCTCTT | TATTAGCGCG | TGAAGGTGGT | AATAACGTTA | 1260 |
| 40 | TTAAATTAAG | TACAGGTGAT | GACGTTGTTG | AATTGTCTTC | TACATCACCA | GAAATTGGTA | 1320 |
| | CTGTAAAAGA | AGAAGTTGAT | GCAAACGATG | TTGAAGGTGG | TAGCCTGAAA | ATTTCATTCA | 1380 |
| | ACTCTAAATA | TATGATGGAT | GCTTTAAAAG | CAATCGATAA | TGATGAGGTT | GAAGTTGAAT | 1440 |
| 45 | TCTTCGGTAC | AATGAAACCA | TTTATTCTAA | AACCAAAAGG | TGACGACTCG | GTAACGCAAT | 1500 |
| | TAATTTTACC | AATCAGAACT | TACTAAAAAT | AAATATAAAT | AAAGGATGAC | GTGATTAATT | 1560 |
| | AAAACGTCAT | CCTTTATTTT | TTGGCAAAAA | TAATTCTAGG | TGCGTATGTA | AAATAAATTT | 1620 |
| 50 | GGCAGCATTT | TAAACAGCAA | ATAAAAGACG | CCAATTAAAT | TTATGACAAA | TGTATCCAAA | 1680 |
| | ATTTAATAAC | тстссттата | TGCCCTTTAA | TTAAAATT | TTAATAGTCA | ATAACAAGTT | 174 |

| AAAAATAAGA | TTATTATT | TATATGTAAA | CGGTTTCTAC | CTCTATTTTA | AATGAAATTT | 1860 |
|------------|------------|------------|------------|------------|------------|------|
| GTGACAAAAA | AAGGTATAAT | ATATTAATGA | CATACAAAGA | AATGGAGTGA | TTATTTTGGT | 1920 |
| TCAAGAAGTT | GTAGTAGAAG | GAGACATTAA | TTTAGGTCAA | TTTCTAAAAA | CAGAAGGGAT | 1980 |
| TATTGAATCT | GGTGGTCAAG | CAAAATGGTT | CTTGCAAGAC | GTTGAAGTAT | TAATTAATGG | 2040 |
| AGTGCGTGAA | ACACGTCGCG | GTAAAAAGTT | AGAACATCAA | GATCGTATAG | ATATCCCAGA | 2100 |
| ATTACCTGAA | GATGCTGGTT | CTTTCTTAAT | CATTCATCAA | GGTGAACAAT | GAAGTTAAAT | 2160 |
| ACACTCCAAT | TAGAAAATTA | TCGTAACTAT | GATGAGGTTA | CGTTGAAATG | TCATCCTGAC | 2220 |
| GTGAATATCC | TCATTGGAGA | AAATGCACAA | GGGAAAGACA | AATTTACTTG | GAATCAATTT | 2280 |
| ATACCTTAGC | TTTAGCAAAA | AGTCATAGAA | CGAGTAATGG | ATAAGGGACT | CCATACCGTT | 2340 |
| TTAATGC | | | | | - | 2347 |

(2) INFORMATION FOR SEQ ID NO: 154:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 13542 base pairs
 - (B) TYPE: nucleic acid (C) STRANDEDNESS: double

 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 154:

| 60 | TGCAACTATC | ATAAAGATTA | GCTCGTCAAG | ATCTGAAATC | TCTATAACTT | ACAAGACGTn |
|-----|------------|------------|------------|------------|------------|------------|
| 120 | TGAAACTCAC | AATCAATGTT | GTCGAAGAAG | AGATGAACAA | ACTGGTTCTT | TCATTCTTAA |
| 180 | CGAAAAAGAA | TATATCTTTA | AGCAATGCAT | CGGCGATGAC | TAACTCGTAT | ATCAATTATT |
| 240 | CAGATAAATA | CTACAATAGA | TTAAACATCA | CGAAGAATAA | GTACATTCGA | CTTGGCGCTC |
| 300 | ATTATTTTGC | CAATGTCTAT | TACAATAACC | TTGGGTCACT | TGATAGGCAT | TCATACGACA |
| 360 | CATTTTCAAG | TGCGTTTTTT | ATCATTGATC | CATTTTCTGA | TCACTAGATT | TTTACGGAGA |
| 420 | AACTCTAAAT | TAATGTTAAT | TTTAAACTAC | TCATTTATTT | GTATTTTTAG | GCTAATTATT |
| 480 | TAATCAGAGT | ATTTACTTTT | GCATATCATC | CGATTTTAAA | ATTAATTTGA | TTGATGTTGA |
| 540 | ATCGTTTGTA | AATATTATGT | CTTCACGTAT | CACGTTATAC | TGATAGATTT | TACATCCAAA |
| 600 | TTGACTACAT | ATTGAAAGGA | CATTTAATTA | TTAATATATA | TAAAAGTCTA | AGCAAATGAC |
| 660 | GTACAAGTAT | AATGGTAAAC | TGAAACAGCA | CACTTGATTT | GCGTTTGTTG | GATACAAGAT |
| 720 | TTCATACTCT | ACAGAAACAT | TAGTCAAATA | AAGTCATTGA | GGAATGGTTA | TTGTTCTGTC |
| 780 | TACAACCAGA | ATTCATGGCA | AAATATTAAA | TTTCACAACA | CAAGACTATT | TGTGAATCCG |

55

5

10

15

20

25

30

35

40

45

| | aGATTTACCT | GTTGTCGCAC | ATAACGCGGC | ATTTGATATG | AACGTCTTAC | ATCAAAGCAT | 900 |
|------|------------|--------------------|------------|----------------|------------|------------------------|------|
| | TCAAAATATT | GGTTTACCAA | CTCCAAATTT | AACTTACTTT | TGTAGTTATC | AACTTGCTAA | 960 |
| 5 | AAGAACCGTT | GATTCGTATC | GATACGGTTT | AAAACATATG | ATGGAGTTTT | ATCAATTAGA | 1020 |
| | TTTTCATGGT | CATCATGATG | CATTGAATGA | TGCCAAAGCA | TGCGCAATGA | TTACTTTTAG | 1080 |
| | GCTACTGAAA | AATTATGAAA | ATTTAACATA | TGTAACTAAT | ATTTATGGTA | AAAATCTAAA | 1140 |
| 10 | AGATAAAGGC | TAGGACTAAA | TAAAATACTC | CCTTCAAAAG | TAAGCATTGT | AAAAATGTAA | 1200 |
| | ACTTTGCAGG | GAGCTTTATT | TTATATAAAG | TCATATATCG | TCATATTTTT | ATAAGTTGAT | 1260 |
| 15 | TGTTCTAAAT | TACCTACAGT | GACACCAATA | AGTCGAATTG | GTACATCAGG | GTCTTTTAAA | 1320 |
| | TCGTTATAAA | GTAAATATGC | AATATTATAA | ATATCTTCTT | CAGAACTAAC | CGAATCTCTT | 1380 |
| | AAACTCATCT | GTTTAGATAG | CGTTTCAAAT | TGATAAGTTT | TAATTTTAAC | CGTTACAGTT | 1440 |
| ?0 . | TTAGCTGACT | TCTGTAATTT | ATTTAGACGT | TCAGCTGTTT | TACCTGNACA | ATTCCCATAC | 1500 |
| | TTTTCTTAAA | ATCTCTTCAT | CATCATTCAC | GTCTGTTGCA | AATGTGCGTT | CAGTCCCTAC | 1560 |
| | TGATTTTCTT | ACTCTTGATG | ATTTCACTTC | ACTATGGTCA | ATACCGCGTG | CCTTGTTATA | 1620 |
| ?5 | TAAACCCCGA | CCTCTTTTTC | CAAACAAACG | TATTAATTCA | AATTCCGTTT | TCTCATATAA | 1680 |
| | ATCTCTACCG | TTAAAAATAC | CATTATCATG | CATTACTTTT | TTGGAAGCTT | TACCTACGCC | 1740 |
| | TGGaAAATCT | CCAATATCCA | ATGTCATCAA | AATATCATGG | aCATTTTGAT | AATCAATCAC | 1800 |
| 30 | AGTCATACCA | TCAGGTTTAT | TCATACCACT | CGCTAATTTA | GCTAAAAATT | TGTTATAAGA | 1860 |
| | AACACCTGCA | GATGCTGTTA | AATGTGTCTG | CTCTAGAATA | TCTTTTCTAA | TATACTGAGC | 1920 |
| ne. | AATTTTCGAA | GCAGGAAGGT | CTGGTCTCAC | TAATTCTGTA | ATATCTAAAT | ACGCTTCATC | 1980 |
| 35 | CAATGACATC | GGTTCTACCT | TATCTGTATA | ACTTCGGAAA | ATAGACATAA | TCTGCGCAGA | 2040 |
| | TGTTTCTCGG | TAAGCACCAA | AATTACTTGT | GACAAAGTAT | CCATTTGGAC | ATAATTTATG | 2100 |
| 10 | CGCTTGTGAC | ATAGGCATTG | CTGAATGGAC | GCCGTATTTT. | CGTGCTTCAT | AGGATGCCGT | 2160 |
| | AGAGACAACA | CCCCTACTGC | TTGCTTTACC | ACCAACAATG | ACTGGTTTCC | CTTTCAATTT | 2220 |
| | GGGGTTATCT | CTCATTTCGA | CTTGTGCAAA | AAAATAGTCC | ATATCTATAT | GAATAATTCG | 2280 |
| 15 | TCTCTCAGTC | AAGTGCTCAC | CTCCCTACTA | ATTTTTACTT | TTATAACGCA | CAAAAATATC | 2340 |
| | TCAACATAAT | TATACGCTGT | GTACGATTTT | TTTACATAAA | TCTTGCACTT | AGCGATAACT | 2400 |
| , | ATATTGAGAT | AACTACAAGT | TGTTATAAAA | TCAATTGCTA | TTTAAGCATG | ATGATGAAGA | 2460 |
| 50 | CGATTGAGTA | AGAAAACATA | GGTAATCTGA | AATAATTCAA | GCAAATTCAT | TTTGTTGGTA | 2520 |
| | TCATCATATT | 7 7 7 7 TOTAL TALE | ATTCACTCCC | CTTTTTTT ATC A | TACAAATAAA | ጥ እ ለተገለ መረጣቸውን | 2580 |

| | AAAGCAATAA | GCGGTATGCA | TACTAAACAT | AAAAATAAGT | GATGAATAAC | CAAATACCTT | 2700 |
|----|-----------------|------------|------------|------------|------------|------------|------|
| | AATTAAAATA | AGCAAGCCAG | TACTTAATAG | GATTAGTGGT | GACAGCATAA | TAATTGAGAA | 2760 |
| 5 | TTGCCATTTG | TTGAAGCAAG | CATCTGCTGT | TTGGAATAAG | ATTCTGTCTT | TTTTTATATT | 2820 |
| | AAACATAGGT | TIGCTATCTT | TTTTAAATAA | AAGAAATAAT | GCTCTATGGA | TAAGTTCATG | 2880 |
| 10 | TAAAATCAAT | AAAATAATGA | ATCCAGCAAA | CCCATATACA | AGATTGATGA | TGATATTTTG | 2940 |
| 70 | ATCGACAACC | GCTGTGACAC | CTAACGCCCA | CTTATACGTA | AATAAAATCA | CGAATAACGC | 3000 |
| | AATAACAAGT | TGCAAGATAA | TAAACCTTCG | CATTTGAAAA | TTATTTGTCG | TTAAATCAAT | 3060 |
| 15 | TTTATGCATT | ACCAACCCTC | CCGATCATGA | CATTCTTATT | CTTCTTTAAA | TATAGTATAC | 3120 |
| • | AATGTCACAT | TTAATTTAAA | AAGTTCATAT | CAAGAAAGTA | AATTGGCTGT | AATAAAATTT | 3180 |
| | TAATATACGA | CTTCTTTCTT | CACTTATTAA | GGCGAAATTT | TATCECAAAT | CATGTGCGCT | 3240 |
| 20 | ATTTCAAATT | GAATAATGCC | ACTGTCTCAA | CATGTGTTGT | TTGTGGAAAC | ATATCTACCG | 3300 |
| | GTGTTACCTC | TTCAAGTTGA | TATTTTTCAG | СТААТААТАА | TGCATCACGT | TGCTGTGTTG | 3360 |
| | CGGGATTACA | TGAAATATAG | ACAATACGCT | TAGGTTCTAA | TGTAAGCAAA | GTCTGAATAA | 3420 |
| 25 | ACGTTTCGTC | ACAGCCCTTT | CTTGGCGGAT | CAACCATTAC | AACATCTGGT | TTAATCCCTT | 3480 |
| | GTGCTTTCCA | TTGTAAAATA | ACTTCTTCAG | CTTTCCCACA | GACAAAAGTT | GTATTATTGC | 3540 |
| | ATTGGTTTAT | AGTCGCATTT | TGTTGTGCGT | CTTCAATTGC | AGAAGGTACT | ACTTCAACAC | 3600 |
| 30 | CGTATACATG | TTTTGCAAGT | GGTGCCATAT | ATAGCCCTAT | TGTTCCAATA | CCACAATAGG | 3660 |
| | TATCTAATAC | AACTTCATTA | CCTGTCAATT | GCGCATACTC | AATTGCTTTA | TTATATATT | 3720 |
| 35 | TCTCTGTTTG | TTCAGAATTA | ATTTGGTAGA | ATGACTGATC | ACTTATTTTA | AATGTACTAT | 3780 |
| | CTGTTAATTG | ATCAATAATT | GTATCTTTAC | CATATAGCGT | TATAGATTGA | CGTCCCATAA | 3840 |
| | TAACĂTTAGA · | GTGGCTATCA | TTAATGTTTT | GTTTAATGCT | TGTCACATTA | GGAAATGCAT | 3900 |
| 40 | CTAATATCTT | CTCAACAACA | GCATTTTTTT | GTGGCCACTT | TTTACCATTA | GTTACAAAAA | 3960 |
| | TAATCATCAT | TTCGTCTGTA | TGATATCCTG | TTCTTACAAC | CAAATGTCTC | ATTAAACCTT | 4020 |
| | TTTTCAATTG | TTCTTGATAA | ATACTTACAT | TTAAATCTTT | TAAAATAGAT | TTAACTTCAT | 4080 |
| 45 | TCATCACTTC | TTGATGTTGT | GAATCTTGTA | TTAAACAACT | TTCCATGTCA | ATAATGTCAT | 4140 |
| | GGCTTCTTTG | ACGATAAAAG | CCCATAATAA | CTTCATTCTG | TTCATTCTTA | CCAACTGGAA | 4200 |
| | TCTGGGACTT | GTTTCGATAT | CTCCAAGGAT | CTGTCATGCC | AACTGTATCG | TTAATCTTAG | 4260 |
| 50 | AATTATCAAA | ATGCGCTTTT | CGCTGAAACA | AATTAATCAC | TTGTTCCTTT | TTCATTTCAA | 4320 |
| | GTTGTGCTTC | GTATGATAAG | TCTTCAACTT | GGCACCCACC | ACBACCTTCA | TAATATATAC | 4390 |

| | AGTTCTTTTT | TACTTTGATA | ATTTTATATT | CAATTTGTTC | AAATTAATTA | GCTTGTGGTA | 4500 |
|----|------------|------------|------------|------------|------------|--------------|------|
| | TGAAAATAGG | AAAGCGATCT | ATTTTTACGA | CACCATGGCC | TTCATGCGTT | AAATCAACAA | 4560 |
| 5 | CTGTTCCCGT | TTTTATGTCA | TTTTTAGCTA | TTGCTTGCAA | AATTTTACCT | CCAAAATGAA | 4620 |
| | CAGGTTAGGA | ACAAAATTAT | GCGCTTCCTA | ACCTGCCATT | ATATATTTCA | CTATTTCTGT | 4680 |
| | TTATTCTTCG | ATTAAATTGT | CATCAACATG | ATCATTATTT | ATTAACTCTT | CATTTACAAT | 4740 |
| 10 | ATCATTAGGT | GCAAAGACAT | CTATATGACG | TTCTAGATTT | AAGAAATTCG | CTGGTAATTT | 4800 |
| | ACCACCATAT | TCTCCATCTA | CATTTAGTTG | TAAGTCTGTG | AATGATGAAA | TATTAATTGC | 4860 |
| 15 | CTTTGCTTTT | TCATAAATAA | CTTTAGGATG | CTTAGTATGT | TCTCCTCTTG | AAGCTAAAGT | 4920 |
| .5 | CATAATATGA | CCAAGTTCTG | CAAGGTTTGA | TTTTTCAACT | ATAATTAACG | TAAAATAGCC | 4980 |
| | GTCATCTAAC | TTAGCGTCCG | GCACTAATTT | TTCAAATCCT | GCCATTGAAT | TTGTTAAACC | 5040 |
| 20 | TAAAAAGAAT | AATAATGCTT | CTCCTTGGAA | AACATTACCA | TCATATTCAA | TTCTTAAATC | 5100 |
| | TACAGCTTTC | ATTTGAGGTA | ACATTTCGAA | ACCTTTGaTG | TAATAAGCAA | ATGGACCAAC | 5160 |
| | AATAGATTTC | AATTTACTCG | GTGTTTCATA | AGAGACTTGC | GTCAATTGTC | CGCCTGCAGC | 5220 |
| 25 | TAAATTAATA | AAGTATCGAT | TATTCATTTT | ACCAATATCT | actttagtåg | AATGACCTTC | 5280 |
| | AATGATGACA | TCAAGTGCCC | CCATGATGTC | ATTAGGTATA | TGCAATGCAC | GTCCAAAGTC | 5340 |
| | ATTAACAGTA | CCCATAGGAA | TGACACCTAG | CTTAGGACGA | TTAGGCTTTT | CTGCGATACC | 5400 |
| 30 | ATTAACTACT | TCATTTAATG | TTCCATCACC | ACCTGCAGCG | ATTAATACAT | CATAATTTTC | 5460 |
| | ATGCATAGCT | CTTTCTGCTT | CAAGTGTGGC | ATCACCTATT | TTCTCGGTTG | CATATGCACT | 5520 |
| | CGTTTCATAT | CCCGCTTTTT | CTAATTTTAT | TAAGGCATCA | GGTAATTCTC | TTTTAAATAG | 5580 |
| 35 | CTCTTTACCT | GATGTCGGGT | TATAAATGAT | TCTAGCACGT | TTCCTCATAT | CTTATCCCTC | 5640 |
| | TACTTAAAAT | TCATATATTT | TAACTTCATC | TTTGTTTCGT | CTAATAGGGA | GTGGGACAGA | 5700 |
| 40 | AATAATATTT | AACAAAATTT | ATTTCGTTCT | ACCCCAACTT | GCATTGTCTG | TAGAATTTCC | 5760 |
| •• | TTTCGAAATT | CTCTATGTTG | GGGCCCCACC | CCAACTTGCA | CATTATTGLA | -AGCTGACAGA- | 5820 |
| | AAGTCAGCTT | CTTTGTTTGG | GGGCCCCGCC | AACTTGCACA | TTATTGTAAG | CTGACAGAAA | 5880 |
| 45 | ATCAGCTTCT | ATGTTGGGGC | CCCACTAGAA | TTGAAAAAAG | CTTGTTACAA | GCGTATTTTC | 5940 |
| | TTTCAGTCAA | CTACAGCCAA | TATAACATTG | TAGTGCCTAG | GACATTGAAT | TTATGACCCA | 6000 |
| | GGCTCAGTCT | TATTTCATCA | TTCTTAATAT | CGTTAAAGAC | CAACTTGTAT | CTTAAACAAA | 6060 |
| 50 | TACTATCTCA | ATATGTACAA | AGCTTGTTAT | TTATTCAGCA | TTTTTTGCCG | TTCTTCATTA | 612 |
| | TALAGCTTCG | TCAGTTATGC | TATTTTACCT | TTAAAATGAT | GTTGTAAATA | TAATGTTGTC | 618 |

| | AACGCATTAA | TAAAATTAAT | ATTTTTACCA | TTAACATGTA | CAATGAATAA | AGTTAAAAGT | 6300 |
|----|----------------|-----------------|------------|------------|------------|------------|------|
| | AATTTGACTT | CTATAGATAT | AAATAAACCC | TCGATTGCAT | CTAAGTCAGC | AATCAAGGGT | 6360 |
| 5 | TTATTTTTTA | AATCTTCATA | GTTTGATGAT | TTAAATTATC | TTTTATCTAA | TTCTTGTTTT | 6420 |
| | AATAGTTGAT | TTACTAATTG | TGGATTAGCT | TGACCTTTAG | ACGCTTTCAT | AATTTGACCA | 6480 |
| 10 | ACTAAGAAGC | CCATAGCTTT | GCCTTTACCA | TTTTTGTAAT | CTTCAACTGA | TTGTTCGTTA | 6540 |
| 10 | TTGTCTAATG | CTTCATTTAC | AAATTTTAGA | AGTGTTGCTT | CATCAGAAAT | TTGAACTAAG | 6600 |
| | CCATTATCTT | CCATAATCTG | TTTAGCATTA | CCACCTTTAG | CTGCTAACTC | TGGGAAGACT | 6660 |
| 15 | TTCTTCGCAA | TTTTACTGCT | CATTGTTCCG | TCTTCGATAA | GTTTAATCAT | ACCTGCTAAA | 6720 |
| | TTTTCTGGTG | TTAATTTAGT | ATCTAATAAT | TCTACTTGAT | TTTTATTTAA | ATATTCGTTT | 6780 |
| | ACGCCACCCA | TTAACCAGTT | AGATGTTAAT | TTAACATCTG | CACCGTGTTC | AATTGTTGAT | 6840 |
| 20 | TCAAAGAAAT | CTGACATTTC | TTTAGTCAAT | GTTAATACGT | GTGCATCGTA | TGCAGGTAAA | 6900 |
| | CCTAATTCAT | TTACATACTT | AGCTTTACGT | TCATCTGGTA | ATTCAGGAAT | TGTCTGACGA | 6960 |
| | ACACGCTCTT | TCCAAGCATC | ATCAATATAT | AAAGGTACAA | TGTCAGGCTC | TGGGAAGTAA | 7020 |
| 25 | CGGTAATCAT | CAGAACCTTC | TTTAACACGC | ATTAAAATTG | TTTTACCTGT | AGATTCATCA | 7080 |
| | AATCGACGTG | TTTCTTGTCC | GATTTCTCCA | CCATTTAACA | ATTCTTCTTC | TTGGCGTTTT | 7140 |
| | TCTTCATATT | CTAAACCTTT | ACGTACATAG | TTAAATGAGT | TTAAGTTTTT | CAATTCGGCT | 7200 |
| 30 | TTAGTACCAA | ATTTTTCTTG | ACCATATGGA | CGTAAAGAGA | TGTTAGCATC | ACAACGTAAA | 7260 |
| | GATCCCTCTT | CCATCTTAAC | GTCTGATACA | CCAGTGTATT | GAATAATTGA | ACGCAATTTT | 7320 |
| 35 | TCTAAATATG | CATATGCTTC | TTTAGGTGAA | CGAATATCTG | GTTCAGATAC | GATTTCAATT | 7380 |
| | AGCGGTGTAC | CTTGACGGTT | CAAGTCAACT | AATGAATACT | CACCTTTATG | TGTTGACTTA | 7440 |
| | CCAGCATCTT | CTTCCATGTG | AAGACGAGTA | ATACCGATTC | GTTTTGTTTC | ACCGTCGACT | 7500 |
| 40 | TCGATÁTCGA | TATATCCATT | TTCACCAATT | GGTTGATCAA | ATTGAGAAAT | TTGATATGCT | 7560 |
| • | TTTGGATTAT | CTGGATAGAA | ATAGTTCTTA | CGGTCAAACT | TAGATTCTGT | TGCGATTTCC | 7620 |
| | ATATTTAGTG | CCATTGCAGC | ACGCATTGCC | CAGTCTACTG | CACGCTTATT | AACAACTGGT | 7680 |
| 15 | AAGACACCTG | GATATGCTAA | GTCGATAACA | TTTGTATTTG | AGTTAGGTTC | TGCTCCAAAA | 7740 |
| | TGCGCTGGTG | ATGGAGAAAA | CATTTTTGAG | TCCGTTTTTA | ACTCTACGTG | AACTTCAAGT | 7800 |
| | CCTATAACTG | TTTCAAAATG | CATGATTTCC | ACTCCTTATA | ATTTTTCATA | AACGTCATGT | 7860 |
| 50 | AAATTGTATT | GTGTTTCATA | TTGATAAGCG | ACACGATATA | ACGTTTTTTC | ATCGAATGGT | 7920 |
| | TTR CCR 2 TC 2 | 1 COCC 1 1 1 CC | a | | | | |

| | GGATCATCAA | TTTCTTCACC | TAAATTAAAC | GCaGTgTnAG | GCGCTGTTGG | ACCAACTACT | 8100 |
|----|------------|------------|---------------------|------------------------|------------|---------------------|------|
| | ACATCATAAT | TTTCGAATAC | TTTATCAAAG | TCATTTTTAA | TCAATGTTCT | AACTITTTGA | 8160 |
| 5 | GATTTTTTAT | AGTAAGCATC | ATAGTAACCT | GAACTTAATG | CAAATGTACC | TAAGAAAATA | 8220 |
| | CGACGTTTTA | CTTCTTTACC | GAAACCTTCA | GATCTTGACA | TTTTATATAA | TTCTTCTAAT | 8280 |
| | GAATGAGCTT | CTTTAGAATG | ATAACCATAA | CGAATTCCGT | CAAAACGAGA | AAGGTTTGAC | 8340 |
| 10 | GAAGCTTCTG | ATGATGCAAT | CACGTAATAT | GATGGAATAC | CAAATTTAGT | ATTTGGCAAT | 8400 |
| | GATACTTCCT | CAACGACAGC | ACCTAAAGAT | TTTAAAGTTT | CTACAGCGTT | TTGAACTGCT | 8460 |
| 15 | TCTTTTACGT | CATCAGCTAC | ACCTTCACCT | AAGTATTCTT | TAGGTAATGC | AACTTTTAAT | 8520 |
| | CCTTTAATAT | CTTTACCAAT | TTCAGATGTA | AAGTCTACAT | CATCAACTGG | TGCACTTGTA | 8580 |
| | GAGTCATTAA | CATCTGCACC | AGAAATAGCT | TCTAATACGA | TTGCATTATC | TTTTACATTT | 8640 |
| 20 | CGAGTCAATG | GACCAATTTG | GTCTAATGAA | GATGCAAAAG | CAACTAATCC | AAATCGAGAT | 8700 |
| | ACACGACCGT | ATGTTGGTTT | CATACCGACA | ACGCCACAAT | ATGCAGCCGG | TTGTCTAATT | 8760 |
| | GAACCACCTG | TGTCTGAACC | TAAGCTAAAT | GGTACTAAGC | CAGCTGCAAc | TGCTGCTGCA | 8820 |
| 25 | GATCCACCTG | ATGAACCACC | TGGCACTGCT | TTATGGTCAA | ATGGGTTAAC | TGTTTTTTTG | 8880 |
| | AAATAAGATG | TTTCTGTTGA | ACCACCCATT | GCAAACTCAT | CCATATTTAA | TTTACCGATT | 8940 |
| | AAAACGGCAT | TTTCATTATG | TAGTTTTTCC | ATTACAGTAG | ATTCGTAAAT | TGGCACAAAA | 9000 |
| 30 | CCTTCTAACA | TTTTACTTGC | ACATGTTGTT | TCTAATCCGT | TTGTAATAAT | GTTATCTTTT | 9060 |
| | ATACCCATTG | GAATACCAAA | TAATTTGCCA | TCCATTTGAT | CTTTTGCTTG | TAATTCATCC | 9120 |
| 25 | AATTCTTGCG | CTTTTTTGAT | TGCATTTTCT | TTATCCAGCG | CTAGAAAAGA | CTTAATTGTT | 9180 |
| 35 | GGATCAGTCT | CTTCAATTGC | ATCATATATA | TCTTTAACAA | CATCAGATGG | TTTGATTTTT | 9240 |
| | TTGTCTTTTA | TTAAAGTTAA | TAAATTCTCA | ACCGATTCGT | AGCGAATGCT | CATCTTACGC | 9300 |
| 40 | GTCCTCCTCA | TTCATGATTG | TAGGCACTTT | AAATTGTCCA | TCTTCTGTTT | CTTTGGCATT | 9360 |
| | TTTCAAAGCT | AATTCTTGTG | GAATACCTTT | AATTGCTTTA | TCTTCACGTA | AAACGTTTTG | 9420 |
| | TAAATCTAAA | ACGTGATATG | TAGGTTCAAC | GCCTTCTGTA | TCAGCGCTAT | CATTTTGTTT | 9480 |
| 45 | TGCAAAATCT | AAAATGCTTT | CTAATGTGTT | GGCCATTTCT | TCCGTTTCTT | CAGGAGAAAT | 9540 |
| | TTGAAGTCTT | GCAAGATTCG | CGATATGCTC | AACTTCTTCA | CGTGTTACTT | TTGTCATTAA | 9600 |
| | TAAAAGCCTC | CTTTAAGTCA | TTCATCACTA | AATTGTATCA | AATTTCCAAT | TAAAAATCTA | 9660 |
| 50 | AGTATTTATG | AGGTGCTACT | TTAATTTCAT | ATAAACTGTA | TAAACATTAT | CATTCGTTTA | 9720 |
| | TC | T | 3 C 3 3 C 3 C T C T | מידיר מידי מידיר מידיר | CACAACCCAA | ተተ ሮል አጥልጥቸል | 9780 |

| | TATATTGGTA | TGCAAGTATT | TCAAAAAGAA | TAAATTTAAT | TTTCCTACTT | TTCTAAACAT | 9900 |
|------|------------|------------|------------|------------|------------|------------|-------|
| | TTATCTTTAT | GTATAATGTT | TTCAAGTAAC | TAAATTATAA | ATTAAATAAA | GGGAGTGTTT | 9960 |
| 5 | ATCATGCTTA | CAATGGGGAC | AGCATTAAGT | CAACAAGTAG | ATGCCAATTG | GCAAACTTAT | 10020 |
| | ATTATGATTG | CCGTCTACTT | CTTGATACTA | ATCGTTATTG | GCTTTTACGG | TTACAAGCAA | 10080 |
| | GCAACTGGTA | ACCTAAGCGA | GTACATGTTA | GGTGGACGTA | tATTGGACCG | TATATTACTG | 10140 |
| 10 | CATTATCAGC | TGGAGCTTCA | GATATGAGTG | GATGGATGAT | TATGGGGCTA | CCTGGTTCTG | 10200 |
| | TCTATAGCAC | TGGTCTATCA | GCTATGTGGA | TTACAATCGG | TTTAACATTA | GGTGCTTATA | 10260 |
| 15 | TAAATTACTT | TGTTGTTGCT | CCTAGACTTC | GTGTTTATAC | CGAATTAGCT | GGAGATGCAA | 10320 |
| ,,, | TTACATTACC | AGATTTCTTT | AAAAATCGTT | TAAACGATAA | AAATAATGTG | TTAAAGATTA | 10380 |
| | TTTCTGGATT | GATTATCGTA | GTATTCTTTA | CATTATATAC | ACATTCTGGT | TTCGTATCTG | 10440 |
| 20 | GTGGTAAACT | ATTTGAAAGT | GCTTTTGGAT | TAGATTATCA | TTTCGGTTTA | ATATTAGTTG | 10500 |
| | CTTTCATTGT | CATTTTCTAT | ACTITCTTTG | GTGGATATTT | AGCTGTATCA | ATTACAGATT | 10560 |
| | TCTTCCAAGG | TGTCATTATG | TTAATTGCGA | TGGTTATGGT | CCCTATTGTT | GCTATGATGA | 10620 |
| 25 | ATTTAAACGG | CTGGGGAACG | TTTCATGATG | TAGCAGCTAT | GAAACCTACA | AATTTAAATT | 10680 |
| | TATTTAAAGG | GTTATCATTT | ATAGGAATTA | TCTCTCTATT | TTCATGGGGA | TTAGGTTATT | 10740 |
| | TCGGTCAACC | TCATATCATT | GTAAGGTTTA | TGTCTATTAA | ATCACACAAG | ATGCTACCTA | 10800 |
| 30 | AAGCTAGACG | TTTAGGTATT | AGCTGGATGG | CTGTTGGTTT | ATTAGGCGCT | GTGGCTGTTG | 10860 |
| | GTTTAACAGG | TATTGCATTC | GTACCTGCTT | ATCATATTAA | ACTAGAAGAT | CCTGAGACAT | 10920 |
| | TATTCATCGT | GATGAGTCAA | GTACTCTTCC | ATCCTCTTGT | AGGTGGTTTC | TTACTTGCTG | 10980 |
| 35 | CGATTCTAGC | TGCAATTATG | AGCACGATTT | CTTCACAATT | ACTTGTAACA | TCTAGTTCAC | 11040 |
| | TAAÇGGAAGA | CTTTTATAAA | TTAATTCGTG | GTGAAGAAAA | AGCTAAAACG | CACCAAAAAG | 11100 |
| 40 | AATTTGTTAT | GATTGGAAGA | TTATCTGTAT | TAGTTGTAGC | AATTGTTGCC | ATCGCGATTG | 11160 |
| . 17 | CATGGAATCC | AAACGACACA | ATTCTAAACT | TAGTAGGTAA | CGCTTGGGCC | GGATTTGGTG | 11220 |
| | CATCGTTCAG | TCCACTTGTG | CTATTTGCAC | TTTACTGGAA | AGGTTTGACA | CGTGCCGGTG | 11280 |
| 45 | CTGTAAGTGG | AATGGTTTCA | GGTGCCTTAG | TCGTTATCGT | TTGGATTGCA | TGGATTAAAC | 11340 |
| | CATTGGCACA | TATCAACGAA | ATATTCGGCT | TATATGAAAT | TATTCCTGGA | TTTATTGTAA | 11400 |
| | GTGTAATCGT | TACATATGTT | GTAAGTAAAC | TTACTAAAAA | ACCTGGTGCA | TTTGTTGAAA | 11460 |
| 50 | CTGACTTAAA | CAAAGTTCGT | GACATCGTTA | GAGAAAAATA | ATTCATAAGT | CTTAACAAAT | 11520 |
| | TAAAAAGGTA | CTAATGTTAA | TCAAAATTAT | GACTAACATT | GGTACCTTTT | TATTATCTTT | 11580 |

| | AATTAAAGCA | CGTGGTTGGT | TACCATCTTT | AATACGAATT | TCATAGTTAT | CGATTTTATC | 11700 |
|----|------------|------------|------------|-----------------------|------------|------------|-------|
| | GAAATATTTA | TTCGCTTGTT | CAGTAACGTA | CTGTGTAATA | CCAATTGTTT | CAGCTTGTCC | 11760 |
| 5 | ATAGTAATCG | ATTGGTAAAT | СТАСТАСТАА | TCGTTGTGGC | TTTTTATCAA | CAAATTTAAC | 11820 |
| ÷ | TTTCCCTACT | GCTTGTGTGA | aattagaaaa | ATATGATTGC | AAATTATCAT | TAAATTGCTT | 11880 |
| | GAAATTATTA | TTTAAATTTT | CATCATAATC | TGCTGCTGTT | GAAGAAGGTA | ATAAAGCTGA | 11940 |
| 10 | TTTTTCATTG | ATATTATGCC | ATTCATTAAG | CTTTGTTTGA | CTCTTTTCTG | CAGTCGCTTG | 12000 |
| | AGTGATAAAT | TCACCTGGTG | TGATTGAATC | TTCACTTGAT | TGCTTATAAA | TTGCAAAATG | 12060 |
| 15 | AATTGGTATA | TCTTTTAAAT | CATCATTTTC | ACGTAACCTT | GATAATATCT | CACTAGCCAT | 12120 |
| | TTGTTTACCT | TGCTTTTTAA | CTCGCTATCA | TCTAGTTTTT | TACTAAAAGT | CGATCCATCT | 12180 |
| | TITTCTTTTT | TATAGTAATA | AACACTATTC | ATAGCTAAAC | CAATCGTCAT | ACCTTTAATA | 12240 |
| 20 | TTCTTACCTT | TTGTATCTCC | ACCACCATAA | AAATCTTGCT | CTAAAATGTT | AGATAAATAG | 12300 |
| | GCTGGTGATT | TTTCTGCAAT | CTTTTCAGGA | TCTGTTTCAC | CTtCGTGTGA | TGGATTAAGT | 12360 |
| | CCTAAATTTT | CATTCGCTTT | CTTGTCTTTT | TTATCTTTTT | CAGACATTTT | ATCGATTTCA | 12420 |
| 25 | CGTTTTGTAT | ACTTAGGATT | TAAATAGGCA | TTAATTGTTT | TCTTGTCCAA | AAATTGACCA | 12480 |
| | TCTTGATACA | AATATTTATC | TGTTGGAAAT | ACTTCTTTAC | TTAAGTTCAA | TAAACCATCT | 12540 |
| | TCAAAGTCGC | CGCCATTATA | ACTATTTGCC | ATGTTATCTT | GTAAAAGTCC | TCTTGCCTGG | 12600 |
| 30 | CTTTCTTTAA | ATGGTAACAA | TGTACGATAG | TTATCACCTT | GTACATTTTT | ATCCGTTGCA | 12660 |
| | ATTTCTTTTA | CTTGATTTGA | ACTATTGTTA | TGTTTTTGAT | TATCTTTTCC | AGCCTGGTCA | 12720 |
| | TCCTTATGGT | TACCACAAGC | AGCGAGTATA | AAGATAGCTG | TAATCAATAA | TACTAATGTA | 12780 |
| 35 | CGCTTCATCG | ACATACCCCT | CTAACTATTT | AATTCATTTT | GCTTATCTAC | AAATTGTTGC | 12840 |
| | TCTGTCCAAA | TTTCAATACC | TAAACTTTGT | GCTTTTGTTA | ATTTTGAACC | TGCATCTTCA | 12900 |
| 40 | CCAGCAATAA | CGACATCTGT | ATTTTTAGTA | ACGCTACTTG | TAACTTTAGC | ACCTTGTGAT | 12960 |
| | GCAAGCCATT | TAGATGCTTC | ATTGCGTGTC | ATTTGATGTA | GCTTACCAGT | CAGTACTATC | 13020 |
| | GTTTTACCAC | TAAATTCAGG | ATGTCCTTCA | ATATCTGATG | TTTTGATACC | TTTATAAATC | 13080 |
| 45 | ATATTAACAT | GTTTATCTTT | TAATTTTTGA | ATTAAAGCAC | GAATATCTTC | ATTTTCTAAA | 13140 |
| | TAAGTAACTA | CAGATTGTGC | TACTTTATCA | CCTATATCAT | GAATTTCTAC | TAATTCCGCT | 13200 |
| | TCAGTTACCG | TTAGTAATCG | ATCTATCGTT | TCATATTTTT | CTGCTAACAC | TTGGCTCGCT | 13260 |
| 50 | TTAACACCTA | AATGCCTAAT | ACCTAGACCA | AATAATAAAT | TTTCTAAAGA | GTTGTCCTTA | 13320 |
| | GCTTGTTGAA | TGGCAGCTAA | TAAATTATCA | PCALALALALACAL | GCCCCATTCT | GTCTAAAGGT | 13380 |

| TAAAGCTGTT GAATAATTTT AGTGCCTAAA CCATCAATAT TCATGGCTTG TCTTGaTACA | 13500 |
|---|-------|
| AAGTGNATCA ATCCLTCAAC AAGTTGTGCT TGGTCATTTT GG | 13542 |
| (2) INFORMATION FOR SEQ ID NO: 155: | ٠ |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1893 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 155:

| 60 | ACTGATTTAC | TAACACATGC | ACATTTATTT | GAATATTTAT | CTCTGATTAC | CAGTAAACAC |
|------|--------------|------------|------------|------------|------------|------------|
| 120 | GGTTATCTAT | TATCTATCTT | AAACATGGTT | TAAAAAGGGT | CACCTTTACG | GACTACTAAA |
| 180 | AAATGAATAT | ATATGTATAA | ATTGCTTAAA | ACGCATAACA | TTnTCATATT | TTATAAATAT |
| 240 | AACATATTtA | CAATTGTTTG | TAATAAGCGT | TTATTAGATT | AACTTGCTAA | ATGTGTAATA |
| 300 | AAAATTCTAT | ATAGAAATTG | ATATTGTCGT | ACAGATACGA | CATTGATATC | ATTAAAATCA |
| 360 | TCGATTAACA | ATGAGAAAAA | AAGTTTCAAC | ACATAATTTT | AAAGTCTTCA | TTTTTAAATG |
| 420 | TTGCTAATCG | TTTACAATTG | CATTTCAAAC | CCTTTTGAGA | GTTGAATATG | AACAACGTCA |
| 480 | ATAAGTGTTT | ACGATTTCTA | AAATAAATCA | CCCTGCTATA | TTTTAGTGAT | ATATATTTGC |
| 540 | TATGATAAGT | GCTGCGTCTC | TTCATCCACT | TTTGCGTTAG | TGTTCATCAA | TGTATTGAAT |
| 600 | TATTATTTTC | TACGTACTTT | TAAACTCACG | CTTTCCCTAA | TCTGCGCCAT | CAATTTATCT |
| 660 | CTAATAAATC | CCATAGCAGT | TGCTTCATCA | TACCTAACTT | CCCACTTTTT | AAGATCGCTG |
| ,720 | GTTCTTTAGT | CTTTCTAAAT | ATAACTATAA | TACCTAAATG | TGGAACATCA | ATCTTTAATC |
| 780 | ATAATGCTCC | GCAAAAGTTA | ACTCATAACC | TATCTGCTGC | ACATTAGCGA | TGTATCATCG |
| 840 | CTTCGCTTTG | ATTGGTTGGC | TTCAAGATCA | TTTCCAAAGT | TTGTGTATCA | TGTTTTTGTT |
| 900 | ACAGCCGTTG | CTTGCTATTG | AACATGACCA | CGACCATTCC | ATTTGACCGC | CATATCTAAC |
| 960 | GTTCAAATGC | CTTGAAATAA | TCTATCATCA | CATCAGTTAA | ATTTTTACTT | TAGAACTTTT |
| 1020 | CTTTATGATT | TCACCATATA | CGCAGTCCAC | CTGCTAATAT | AAAGCATCAC | TTTAGTTAAT |
| 1080 | CATGAATAAG | GGTAGGTCAT | ATCCATCGCT | AATCATCATT | CCTCGTCGAT | TGTTAATTTT |
| 1140 | CATACTCGGT | ATACCTAACT | TGCGCTCTTC | CTAGTGCAAT | TGAATCATTT | TGAATATGTA |
| 1200 | CTCCAGCATT | ATGCGTTTAC | AACTGGTCGG | GTAATAACAG | TCTAAAGTGA | ATTTAGTGAA |
| 1260 | TTATCGCAAC . | ACTGATTTAT | AGTATCCATT | CTTCTAGCTG | AACATACTTT | TAATGAATAC |

55

5

10

15

20

25

30

35

40

45

| CATCCTCAGC | TTCTTCTTTT | ATTAAGTCAT | TCACCTTTTT | TTCGGCATTT | TTTAAAGTTG | 1380 |
|------------|------------|------------|------------|------------|------------|------|
| TGTCACAAGC | TGCTGATAGT | TTCATACCAC | GTTGATATAA | ATCTAATGAT | TCCTCTAAAG | 1440 |
| ATACTGTTTC | ATTATCTAAT | TTTTGAACAA | TTTGCTCTAA | TTCTTGCATC | ATTTCTTCAA | 1500 |
| AACTTTGCGT | TTCTTTAGTC | ATTATTACAC | CTTACTTTCG | TAACTTTTGC | ATCTACTAAG | 1560 |
| CCATCTTTCA | TTGTTAACGT | CAATTGATCA | TTTTCTGTTA | AATCTTTAGT | ACTCGTAATG | 1620 |
| ACTTCGTCTT | TTTTATTAAC | AATTGCATAT | CCACGCAACA | TTGTATTAGT | TGGACTTAAA | 1680 |
| TTGTTTAAGT | TTTCTACTTT | ATTTTTCAAA | TCATTTTTAT | AACTTAATAT | CTTAGAATTC | 1740 |
| AATAATTTAA | CAAGTTGGTT | TGTCAATTGA | AGATTATnTT | GTTGTTCTTG | ATTAACACTA | 1800 |
| CTTAGTAATG | CTTTTAAATn | ATAACGTTGG | TGCAACAGCA | TTAAATCGAG | GCCCCGGTGG | 1860 |
| TCCAAAGTTG | CCCGAATTnG | TGGTTTCAGG | ccc | | | 1893 |
| | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 156:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 821 base pairs (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 156:

| .c 60 | TAATATGTTG | ATGGTGATTG | AGAGAAAAAC | ATATTCAATT | CCTTCACTTA | AAAATATATT |
|---------------|------------|------------|------------|------------|------------|------------|
| IA 120 | TAAGGTTGGA | AATCGTAAGC | TTTATTTCTG | TAATACTTTT | CTGGGTGTTT | TGCAATATTT |
| T 180 | TAAAAAACTT | TTTCCTTCAT | AATAATTTTA | GTAAAGGGAA | AATAACTCTA | AATTTATAAT |
| AT 240 | CTCTAAGTAT | TCAATAATGG | TGGCGGTCTT | ATTTTTTATT | ATATTTTATT | TTGTTTCACA |
| GC 300 | TAATGACTGC | GGTTGGTTGG | AGGCTTTATT | AAAATATAAT | TTAAATCCTC | GTTATTAACT |
| A 360 | TATTATCTAA | AAAATTTATG | TATTGACAAA | TGTCATCGAT | CTGTTAAACA | AGGTTTCTTT |
| A 420 | ATTTCGGCAA | TGCAGGTTCT | CTGGATTTAC | GATGGTTTAG | GTGGAAAAAT | AACTAACACG |
| A 480 | AACGATTGGA | AAAAGATAAA | AATTTGGAAT | CAAAAAAATG | TTGGACCAAT | TACTTGTATA |
| AG 540 | AAAACAGTAG | ATCTGCAGAA | CTGTAGAAAA | GGTATAGATG | ACTAGACGTT | TAGGACATAA |
| A 600 | TAAGCCCTAA | TCTAATCATA | AAAAAGTATT | GTGAAGCTTC | AAATGTCATG | ATGGTGTTGA |
| ra 660 | ACCCAAAATA | ATCAATAAAA | ATCTAAAATC | TGCTATGCGA | TGGTAAATGT | GAAATGGAGC |
| AG 720 | CTTGGATTAG | ATAATAAGCA | AATGGTCGAT | ATGAATACTT | AAATATAATG | TAGAATTATT |
| A 780 | TTAGTAGAGA | TACGTCAAAA | CCCaAAAGaA | ATTGGTTCAT | CCTTTTATTA | TTTATTTTTT |

55 .

5

10

15

20

25

30

35

40

45

(2) INFORMATION FOR SEQ ID NO: 157:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2343 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

10

15

20

25

30

35

40

45

50

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 157:

| AGTAAGATAA | TTTTCAATTA | GAAAATATCT | TACTGCTGTT | CICIATITAT | ACAATACTTC | 60 |
|------------|------------|------------|------------|------------|------------|------|
| GTATTGAATG | GCTTCGCTTT | CCTAGGGTGC | CGTCTCAGCC | TTGGTCTTCG | ACTGGCACTG | 120 |
| CTCCCTCAGG | AGTCTCGCCA | TTAATACTAC | GTATTAACAT | GTAATTTTAC | TTTGAAATAC | 180 |
| TTTTAAAAAA | TAAGACACTT | TGCCCAACTT | GCACATAAAT | GTAAAATTCA | ATAAAATGAA | 240 |
| TTTTCTGTGT | TGGGTCCCTT | CTTATAATTT | AATAAATACC | ACTAAACTAA | ATTAACGAGG | 300 |
| TGCCTTATGT | АТАААААТТА | TAACATGCCC | CAACTACACT | ACCAATAGAA | ACTTCTGTTA | 360 |
| GAATCCCTCA | AAATGATATT | TCACGATATG | TTAATGAAAT | TGTTGAAACr | ATACCTGATA | 420 |
| GCGAATTCGA | TGAATTCAGA | CATCATCGTG | GCGCAACATC | CTATCATCCA | AAAATGATGT | 480 |
| TAAAAATCAT | CTTATATGCA | TATACTCAAT | CIGTITAATT | ATGTTCAAAG | CATTAAGGTA | 540 |
| ACAAGACAAT | ATCTAAGATA | TCAAAGATAG | AAATTTTTTG | ACGTTGTTGC | TGATTGTAAA | 600 |
| CATAACCATC | AATTTCATAA | TTAATAGCAT | CAATACGATA | AATGGTTAAG | CGTACTGAAT | 660 |
| CTACAAAGCC | ATTATTATAA | AATTTAACTT | CTACAGGTTG | GGCATATTGT | AGCGCCTCGT | 720 |
| GTAGCCGAAT | GTTTAGCTCA | GCCAATTGAT | CATCTGATAA | TACAGGACGT | GTAATTTTGT | 780 |
| TTTGGTCGAT | AATGTATTGT | TGAATCGTTT | CGAATTGTTC | GGGTAATGTT | GCAAAAGGAG | 840 |
| CCCATTTAAT | CATGCCTCTT | CCCATAGGTA | TATTGTTATC | TAGTAATTCT | CTTGGAACGT | 900 |
| TACGATAATC | AGTTTCTTCT | TCATAACTTG | TCATCCTTAA | TTCACCCCAA | TCTGATAATT | 960 |
| ACATTATACG | AACATGTGTT | CTATTTTGCA | ACAAAAATTT | TGTGGAAGCA | TAAACGCGTT | 1020 |
| AATAATTAAT | GCTCGTGtAA | GTAAAAAAGA | GGGATTAATT | AAAATCGAAT | AATGACATAT | 1080 |
| CACaGCAAAT | AGTTCTTTTA | AAGTAGTTAA | ATAGTTTTAG | CTTTAAGGAA | aTGATAAaTG | 1140 |
| ATTGTWAATT | CTAGCTAAAA | TTTAATAAAA | TGAAAATAAG | ACTAACATGG | AGGGGTAAAA | 1200 |
| GTAATGACAA | ATGGATATAT | TGGTTCTTAC | ACTAAAAAGA | ATGGTAAAGG | GATTTATCGT | 1260 |
| TTTGAATTAA | ACGAAAATCA | GTCACGTATT | GATTTATTAG | AAACAGGATT | TGAATTAGAA | 1320 |
| GCGTCTACAT | ATTTGGTGCG | TAATAATGAA | GTTTTATATG | GAATCAACAA | AGAAGGAGAA | 1380 |

| | TGTTTGTCTT | CAAAAGCTGG | TACAGGTTGT | TATGTATCGA | TTTCAGAAGA | TAAACGATAT | 1500 |
|----|------------|------------|------------|------------|------------|------------|------|
| | TTATTTGAAG | CGGTATATGG | TGCTGGCATC | ATACGTATGT | ATGAATTAAA | TACGCACACA | 1560 |
| 5 | GGTGAAATTA | TACGTCTAAT | TCAAGAACTT | GCACATGATT | TTCCAACAGG | TACACATGAA | 1620 |
| | AGACAAGATC | ATCCACACGC | ACATTATATT | AATCAAACTC | CAGATGGTAA | GTACGTTGCA | 1680 |
| | GTAACAGATT | TAGGTGCTGA | TCGTATCGTT | ACTTATAAAT | TTGATGACAA | CGGGTTTGAA | 1740 |
| 10 | TTTTATAAAG | AATCTTTATT | TAAAGATAGT | GATGGGACAA | GACATATTGA | ATTTCATGAT | 1800 |
| | AATGGAAAAT | TTGCTTATGT | CGTACACGAA | TTATCAAATA | CTGTGAGTGT | TGCAGAATAT | 1860 |
| 15 | AATGACGGTA | aatttgaaga | GCTCGAGCGT | CATTTAACAA | TTCCTGAAAA | CTTTGATGGA | 1920 |
| | GATACTAAAC | TTGCAGCAGT | GCGTTTATCT | CATGATCAAC | AATTCTTATA | TGTATCTAAT | 1980 |
| | AGAGGGCATG | ATAGCATTGC | AATTTTTAAA | GTTCTTGATA | ATGGTCAACA | CTTAGAACTA | 2040 |
| 20 | GTAACAaTTA | CTGAAaGTGG | TGGTCAATTC | CCAAGAGATT | TTAATATTGC | CTCATCAGAT | 2100 |
| | GACCYTTTAG | TTTgTGCTCA | kGaGCaAGGA | GATTCAGTTG | TAACTGTTTT | CGAAAGAAAT | 2160 |
| | AAAGAAACAG | GTAAAATTAC | GCTATGTGAT | AACACTCGTG | TAGCATCTGA | AGGTGTATGT | 2220 |
| 25 | GTCATATTTT | AATCTTTAAT | TAATCATGAT | AAAAAGAAAA | CCATGTTTCC | AAAAATTTG | 2280 |
| | TGTATACCTT | GAAATTTATT | GnTTTCCAGn | ACATCAATTA | TGGGAAGCAT | GGnTTATTTT | 2340 |
| | TGT | | | | | | 2343 |

(2) INFORMATION FOR SEQ ID NO: 158:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4837 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 158:

| 60 | AGGTATCATT | GTTTCTTCCA | GTAGTCGAAG | TTCTGGTGCA | TTGGTATCGC | AAATTGCCAG |
|-----|------------|------------|------------|------------|------------|------------|
| 120 | TGAATCTGGA | AAATTACTGT | TTGTTATACA | GATGGCAGTA | ATATCGTTAT | CCGATTGGCT |
| 180 | TATTCAAGTT | AAGACCAACG | AATATTTCAC | TAGTATTACA | CAATTCAAGA | CAATTTTTAA |
| 240 | TGGTGTACCA | CAGCAGGATT | TTAGAAGGTG | CAACGCATTT | GATTTGCATT | TTACTTATTG |
| 300 | AGCTGCGATG | ATCCATTAAA | TTAGGATTTA | ATTAACACAA | GTGCACTTTT | ATTGCAATTT |
| 360 | CCCTGTAGGT | CGATTGGTAT | GCTTTTGGTG | AGCGTCTGGT | TCGCAAATGC | TTATGTTTAG |
| 420 | TCAATCAGCA | TAGGTGTTTC | GTTTCAGTAT | ACCTGGAGAT | CGTTGAAATT | GTTGTAGAAA |

55

30

35

40

45

| | GGTTTTAGAG | GTGTTAAAGA | AACATTACCA | GCAATTTTAG | TAGTTTCAAT | CACTTATACA | 540 |
|-----------|---------------|------------|---------------|-------------|---------------|------------|------|
| | CTTACTCAAG | GATTATTAAC | TGTATTCAGT | GGACCTGAAT | TAGCAGATAT | TATTCCACCG | 600 |
| 5 | TTATTAACAA | TGTTAGCATT | AGCAGTATTT | TCTAAAAAAT | TCCAACCAAA | ACACATTTAT | 660 |
| • | CGTGTTAATA | AAGATGAAGA | AATTGAACCT | GCAAAAGCAC | ATTCTGCAAA | AGCAGTATTA | 720 |
| | CATGCATGGA | GCCCATTCAT | TGTATTAACA | GTCATTGTAA | TGATTTGGAG | TGCGCCATTC | 780 |
| 10 | TTTAAAAACT | TATTCTTACC | AAATGGTGCT | TTATCATCAT | TAGTATTTAA | ATTCAACTTA | 840 |
| • | CCTGGaACAA | TCAGCGAAGT | TACGCATAAA | CCATTAGTAT | TGACTTTAAA | TATTATTGGA | 900 |
| 15 | CAAACAGGTA | CAGCTATTTT | ATTAACTATT | ATTATTACAA | TTTTAATGTC | TAAAAAGGTT | 960 |
| | AACTTTAAAG | ATGCAGGTAG | ATTATTCGGC | GTTACATTTA | AAGAGTTGTG | GTTACCAGTT | 1020 |
| | CTTACAATTT | GTTTCATCTT | AGCAATTTCT | AAAATCACAA | CTTATGGTGG | TTTAAGTGCA | 1080 |
| 20 | GCAATGGGTC | AAGGTATTGC | TAAAGCAGGT | AATGTCTTCC | CAGTTCTATC | ACCAATTTTA | 1140 |
| • | GGTTGGATAG | GTGTGTTTAT | GACAGGATCA | GTTGTAAATA | ACAACTCATT | ATTTGCACCA | 1200 |
| | ATTCAAGCTT | CTGTTGCACA | ACAAATTGGA | ACAAGTGGTT | CACTTCTTGT | ATCTGCTAAT | 1260 |
| 25 | ACAGITGGIG | GTGTAGCGGC | AAAATTGATT | TCACCACAAT | CAATTGCAAT | TGCAACTGCA | 1320 |
| | GCAGTAAAAC | AAGTTGGTAA | GGAATCAGAA | TTATTAAAAA | TGACATTGAA | ATACAGTGTA | 1380 |
| | TGTTTACTAA | TATTCATCTG | TATTTGGACT | TTCATCTTGT | CATTATTATA | AAAAAACGTA | 1440 |
| 30 | TTTCAAAATA | TAAATATACA | GAAGGTGAGA | TGTTTTCTAA | CATCTCATCT | TTTTTTTATG | 1500 |
| | GATCATTAAT | GAAAGAAGTT | TGACATTATA | ATAATGGTAG | CGCTTTATGT | TAAAATGAAT | 1560 |
| | AGTGAGTAAT | CAGCAATCAA | ATTAAATTGG | TTGATAGCTG | TTAAGGTTTG | TGGTTTTGTC | 1620 |
| 35 | TTTGTGCTAT | CGCnCATAAA | GTATATAATT | AAAGTAGTTT | CGTTATTATA | AAATATTAAT | 1680 |
| | ATAÇATAGTA | GATAGTAATA | GAGCATCACC | ATGGGAACCT | ATTGAGACAC | TTATTGATTT | 1740 |
| 40 | AAAGTGGTAT | TAATATGTCG | TATTTCTCGA | ACGTTCCATT | ATTCATTTTA | AAAAGGGGGA | 1800 |
| | CTGTATTTGT | TATGACAACA | CAACATAGCA | AAACAGATGT | CATCTTAATT | GGTGGCGGTA | 1860 |
| | TTATGAGTGC | aCATTAGGAA | CATTACTTAA | AGAATTATCA | CCTGAGAAAA | ATATTAAAGT | 1920 |
| 45 | GTTTGAAAAA | TTAGCACAAC | CTGGCGAAGA | GAGTTCAAAT | GTATGGAATA | ATGCCGGTAC | 1980 |
| | AGGGCATTCA | GCACTTTGCG | AGTTGAACTA | TACAAAAGAA | GGTAAGGATG | GCACAGTTGA | 2040 |
| | TTGTAGTAAA | GCAATTAAGA | TAAATGAGCA | GTACCAAATT | TCAAAACAGT | TTTGGGCATA | 2100 |
| 50 | TTTAGTTAAA | ACAGGACAAT | TAGATAACCC | AGATCGCTTT | ATTCAAGCGG | TGCCACACAT | 2160 |
| | CVCALALALACAC | ATTCCCCAAC | ATTA ATTOTACO | *********** | a concorrence | CAACCTTAAA | 2220 |

| | GGTACCGTTA | ATGATTGAAG | GTCGTAAGTC | TGATGAACCA | ATTGCTTTAA | CTTATGATGA | 2340 |
|----|------------|------------|------------|------------|------------|------------|------|
| | AACTGGTACa | gATGTTAACT | TTGGTGCGTT | AACTGCAAAG | TTATTTGATA | ATTTAGAGCA | 2400 |
| 5 | ACGTGGTGTG | GGAATTCAAT | ATAAGCAGAA | TGTATTAGAC | ATCAAGAAAC | AGAAATCTGG | 2460 |
| | GGTATGGCTA | GTTAAAGTTA | AAGATTTAGA | AACTAATGAA | ACGACAACAT | ATGAATCTGA | 2520 |
| | TTTTGTATTT | ATTGGTGCTG | GCGGTGCGAG | TTTACCATTA | CTCCAAAAGA | CTGGGATTAA | 2580 |
| 10 | ACAATCAAAA | CATATTGGTG | GTTTCCCGGT | AAGTGGATTA | TTCCTGCGCT | GTACAAATCA | 2640 |
| | AGAAGTGATT | GATCGTCATC | ATGCTAAAGT | GTACGGAAAA | GCAGCAGTGG | GTGCGCCACC | 2700 |
| 15 | AATGTCAGTG | CCGCACTTAG | ATACACGTTT | TGTAGACGGC | AAGCGTTCAT | TGTTATTTGG | 2760 |
| | TCCATTTGCA | GGTTTCTCAC | CTAAATTTTT | AAAAACAGGT | TCACATATGG | ATTTAATTAA | 2820 |
| | ATCGGTTAAA | CCAAATAATA | TCGTGACGAT | GTTATCTGCA | GGTATCAAAG | AAATGAGTCT | 2880 |
| 20 | TACGAAGTAT | TTAGTGTCAC | AATTGATGTT | ATCTAATGAT | GAGCGTATGG | ATGATTTAAG | 2940 |
| | AGTCTTTTTC | CCAAATGCTA | AAAATGAAGA | TTGGGAAGTG | ATTACAGCAG | GGCAACGTGT | 3000 |
| | CCAAGTAATC | AAGGATACTG | AGGATTCTAA | AGGTAACTTA | CAATTTGGTA | CTGAAGTTAT | 3060 |
| 25 | TACGTCAGAT | GATGGCACAT | TAGCTGCATT | ACTTGGTGCA | TCACCTGGTG | CGTCAACAGC | 3120 |
| | TGTAGATATT | ATGTTTGATG | TTTTACAGAG | ATGCTATCGT | GATGAATTCA | AAGGATGGGA | 3180 |
| | ACCAAAGATT | AAAGAAATGG | TGCCGTCATT | TGGTTATCGC | tTAACAGATC | ATGAGGATTT | 3240 |
| 30 | ATATCATAAA | ATTAATGAAG | AAGTAACTAA | GTATTTACAA | GTTAAATAAT | AAACGAAACG | 3300 |
| | GTAATGTCTT | TTTTAATGTG | ATAGACATTA | CCGTTTTTTA | GTGGTTAATA | AAAATCATTT | 3360 |
| 95 | TAATTGTTTC | AGTTGCTTGT | TAATAGTGTC | TACGTAGTTC | TTGTTTTTAA | AGAATTGAAT | 3420 |
| 35 | TATCCAAATT | AATACATAAA | CCACAATGAA | GATAATTGTG | AATATGATTA | GATAATGCAC | 3480 |
| | TGTTAGTGGA | AACCAACCGG | CAAGCATTGC | TAAAGGCAAG | AATCCGACAT | ACGTTGTTAT | 3540 |
| 40 | GAAATGCATT | ATAGTTGCTT | TAGTAATGCT | CCAATCTGTG | TATTTAAAGA | TAAAATCTCC | 3600 |
| | AAGGAAAAAG | ACGACGCCTA | TGAGTAACCA | TAAAATGATA | GAAATCAACA | TTACGGTAGT | 3660 |
| | TTCTGTGAAA | TGCGTATAAT | ACAATATGCC | AATAGTTGAT | TGTGGGTTCA | GTGGATAATA | 3720 |
| 45 | TTTGCCGTCT | GCAAATAACA | TACTAAAGAA | CAGTGAAAGG | GACAAACCAA | TGATTAAGCT | 3780 |
| | AATAAATAAT | GAGTTTTTCA | AATTTTTCAT | ATTGATAAGC | GCTCCTTTAT | AGATTTTAAA | 3840 |
| | TAACGTCTAG | AAGAATAGGT | GTAGTGTGCA | TCTTTAAGAT | ACATACGTAT | AAGTCCATTT | 3900 |
| 50 | GGCTCTAATA | ATAATTTTTC | AATGTAATAC | TTGTTGACGA | TTTCTGATTT | GGAAATGCGA | 3960 |
| | ATGAAATGTT | GTGGTAACTG | TTTTTCTAGT | TCATAAAGTC | GTAATTTTAG | TTTGAATTTT | 4020 |

- 3. - 2.

| | | | | | | ~* | |
|------------|-------------|---|--|------------|------------|------------|------|
| | ACATTAATGA | TATGGATTTC | TTTGTCTATG | TATCCGACTA | ATGTATGTGA | TTTGTCTAAA | 4140 |
| | TCATTGACTG | CATTAATAAT | ACTTTGAACG | TTATCATTCA | TTTTAGGTGC | ATGTATATCA | 4200 |
| 5 | ATATAAGATT | CCGTCTCATT | TGCATTGATA | AATAAATTGA | GTTTCATCAT | AGGTTAATGC | 4260 |
| | CTCCTTCAAA | ATTATTAAAC | CATAAATGAC | CATCGATATA | TTTAAATTTT | GTTGAATGGT | 4320 |
| | AGAAATTAAA | TGTTAAGTGG | CTAGAAAGCG | СТААТСААТА | TAAAAGATAC | CTCCTGAAAT | 4380 |
| 10 | AAAAACAGAA | ATGTTTTTC | AGGAGGTAGA | GATTAAAGTG | AATTATTTGG | CAGTGTAATA | 4440 |
| | GTAAAGGTGG | TTACATACTC | GTTACTTTGT | GTGAATTGGA | TTGTACCATG | ATGCAATTCA | 4500 |
| 15 | ATGATGGATT | TTGTAATTGC | AAGACCTAAA | CCATTGCTAT | TATCATGTTT | GCTCACTTTA | 4560 |
| | TAAAAACGTT | CAAATAAACG | TGCTTCAGCT | TGTGGACTAA | TTGGTGAACC | ATCATTACTT | 4620 |
| | ATTGTGAAAA | TGATATTGTT | GTGACTATGT | TGCAAAGCGA | TGTCAATGGC | ACCACCAACA | 4680 |
| 20 | TCTGTATACT | TAATAGCATT | TATTAATAAA | TTACTCAATG | CTTGATGTAA | CAAACGTTGA | 4740 |
| | TTTCCTAGGA | AATTGATGAT | TCTAGGTCAG | CTAAnATGAT | TAACGACTTT | TCATCAGCAG | 4800 |
| | CANATTGTTC | ATGTCGAATG | ATATCTTAA | TGAGCTG | | | 4837 |
| 25 | (2) INFORMA | ATION FOR SE | Q ID NO: 15 | 59: | | | |
| 3 <i>0</i> | (| QUENCE CHAR A) LENGTH: (B) TYPE: nu (C) STRANDED (D) TOPOLOGY | 1600 base p cleic acid NESS: doubl | pairs | | | |
| | | | | • | | | |

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 159:

| 35 | ACAATTATTG | GATTATTATC | AAGCAACGTT | AATGGATGAC | TTCCACTTAC | AACAGAAATG | 60 |
|----|------------|------------|------------|------------|------------|------------|-----|
| | CCCATAGATT | CTAAATCTtT | TGCATGAGCA | TCTTGTGATA | AGTCTTTTCC | ATCATTGACA | 120 |
| 40 | GTTACATTCG | CACCTAATTT | ACTTAATAAT | TTAGCTGCTT | CATAACCACT | TTTTGCCAAA | 180 |
| | CCGACAACTA | ATACATTTTT | ATTTTCTAAC | CCTGTATAAT | TAAGCATCTT | AATGCACTCC | 240 |
| | AATCCATAAA | CCGATTAAAC | CTGAAATCAG | ACCAACAGCC | CAAAATACTG | TAACTACTTT | 300 |
| 45 | CCATTCGCTC | CATCCTATCA | ATTCAAAATG | ATGATGAATC | GGACTCATTT | TAAATATACG | 360 |
| | CTTTCCAGTC | AATTTAAAGC | TAGCGACTTG | TAACATAACA | GATAATGTTT | CAATTACGAA | 420 |
| | TACTAAACCT | ATAAAAATTA | ATGATAATTC | CTGATTAAGC | ATGATTGAAA | TGGTAGCAAA | 480 |
| 50 | TATACCACCT | AAAGCTAAGC | TACCTGTATC | TCCCATAAAC | ACTTTAGCAG | GGTTAATGTT | 540 |
| | ATATGGTAAA | AATCCTAAAA | GTGCAAACAA | CATAATGATA | CAGAAAATAC | CAATTGCCGT | 600 |

| TGCTAATCCA | TCTAAACCAT | CTGTTAAATT | TACTGCATTA | GAAAAACCTA | CTTGCCAAAA | 720 |
|------------|------------|------------|------------|------------|------------|------|
| AACAATGAAA | ATAACATATG | CAAATGATAG | TGGGATTGCT | ACATTCGTAA | ATGGAATATG | 780 |
| TATGCTCGTA | GAAAAATTCA | CCAAATGAAA | CACATTACTT | AAAACAAAGA | ATATAATCGC | 840 |
| AATACCAATT | TGCGCCAAAA | ACTTCTGTTT | ACTTGTTAAA | CCTTGGTTAT | TCTTTTTAAC | 900 |
| AACAATAATA | TAATCATCTA | TAAAACCAAT | TAACCCAAAA | CCAATCGTCA | САААТААТАА | 960 |
| CAGTATGATT | GGATTAGCTT | GATCTACAAA | TATAATAGCC | ACCAAAGACG | TTATCACAAT | 1020 |
| ACTTAATAGA | AATGTTAGTC | CACCCATCGT | TGGTGTACCA | GTCTTCTTCA | TATGGCTTTG | 1080 |
| TGGACCTTCT | TCTCGAATAC | TTTGACCAAA | TTTCATCCTT | TTTAATGTAG | GTATTAAAAC | 1140 |
| AGGTACCAAA | ACAAATGTAA | TCACTAGCGC | TAATAACGCA | TATACAAAAA | TCATAACTAT | 1200 |
| CTCCTCTTCT | TAATCCAGAC | TTTTTTAACC | ACTAATATAT | TATCAATTTT | TCAATTAAAT | 1260 |
| AAACAAAGTT | GTAATCAAAA | TTTATAATTT | TTCTTTTTTA | CGGCATAAGA | GGCCAGTATA | 1320 |
| AAAAGTTTGC | CTATAACAAA | CAAGTTAATC | TGACCTCGTC | TACCTTAAAA | TTCTCTATCA | 1380 |
| ACACTTATTT | ATAAAGATTA | AATGAAGATG | TTGTTTTCTA | TCACAGCATT | ACTTTAGTAA | 1440 |
| AAACAAATAG | TGACAATACA | TCCTAATTTA | ATGTAGCCAT | TCTTGTTAGT | CCGACTTATC | 1500 |
| CTTGTCAGTT | TTACTGTCAG | ATTTCnTCTT | ATCATCTGAA | TTTGAATCAG | AATTATTCGT | 1560 |
| CGAATTGCTG | TCTACATTCT | CTGGATGGAA | AATTCTACGT | | | 1600 |

(2) INFORMATION FOR SEQ ID NO: 160:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1186 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

:

5

10

15

20

25

30

35

40

45

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 160:

ATTGCCTTTG TTTTAATTTT AAATCAAAAT MGCCTATGAA AGATTTAAAT CAATTAATTT 60
CTATAATATT ATCATTTTTA AAGCATATCA TTGTTTAGTT TTTTTATAAT TGGATAAATA 120
CTAATAGTTA CTTTATAAAA CATTACATAG AGAAAGGTTA AGGAGTGCAC ATGTCGAAAA 180
AGGATCACTC TTCTTCAAAA TACCTTAATT CTGTTAAGGA AGCGCAAGAG GAGTCAAAAA 240
AGAAAAATAA AAGTAATCCC AAAATTGATG TTGATCGTAC ATATATTGAA CCTCAACAAT 300
TCCAATCTAA GAAACCTAAA AAAGATGATC AGGTTTTCTT CTTATCAAGA TTAAATAAAC 360
CTGCAAAATA TAAGAAAGAC TCTAATTTCT TATCATATCT CATCTATCGC ATAGGAAAAG 420

2.72

| | TGTTGCTTTT | CCTATTAACA | TTATTACCAT | TTTTCAATAT | TAAGCAGAGT | CAAATTACTA | 540 |
|---|-------------|-------------|-------------|------------|------------|------------|------|
| | ATATGTTAAG | CAATGCACCC | GCTGAAACAT | CTACTCTAAT | TAAGAGTGTA | ATTGGTGATA | 600 |
| | TAACTCAAAA | CTCCAGTGGT | GGCTTATTAT | CTATCGGTTT | GATTTTAGCA | ATTTGGTCAG | 660 |
| | CTTCAAATGG | AATGACTGCA | ATTATGAATT | CTTTCAATGT | TGCTTACGAT | GTAGAAGATA | 720 |
| | GCCGTAATGG | AATCGTATTA | AAACTACTAA | GTGTTGTCTT | CACTGTAGTT | ATGGGCGTTG | 780 |
| | TGTTTGTAGT | TGCTCTAGCA | TTACCAACGC | TTGGTTCTGT | AATTAGTCAT | TTCCTATTCG | 840 |
| • | GTCCACTTGG | aTTTGACGAA | CAAGTGAAAT | GGATTTTTAA | CCTTATTAGA | ATTGTGTTAC | 900 |
| | CAATCATTAT | TATATTTATC | ATATTTATCG | TGTTATATTC | GGTTGCACCT | AACGTTAAAA | 960 |
| | CGAAGCTTAA | GTCAGTATTA | CCAGGTGCAG | TATTTACTTC | AATTATTTGG | TTAGCTGGTT | 1020 |
| | CATTTGGTTT | TGGTTGGTAT | ATTTCAAATT | TTGGTAACTA | TTCTAAAACA | TATGGCAGTA | 1080 |
| | TCGCGGGTAT | CATCATTTTG | TTACTATGGT | TATATATCAC | AAGTTTTATT | ATAATTGTCG | 1140 |
| | GnGCTGAAAT | CAATGCAATC | ATTCATCAGC | GTAGTGTAAT | TAAAGG | | 1186 |
| | (2) THEODAY | TTON FOR CE | O TO NO. 10 | | • | ₹ | |

(2) INFORMATION FOR SEQ ID NO: 161:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7872 base pairs
- (B) TYPE: nucleic acid (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

30

35

40

45

50

5

10

15

20

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 161:

| TCTTGAGCCA | TCTTTTGAGC | TAACTGACTA | GATTGATACC | CAAAAATCAT | AGTTACCAAC | 60 |
|----------------|-------------|------------|------------|------------|------------|-----|
| ATAAACTTTA | ATTTTACCGA. | AGTCTAAATC | AGCGATATGA | GTACATACAT | TATTTAAGAA | 120 |
| ATGACGGTCA | TGCGATACTA | CGATAACAGT | ATTATCAAAG | TTAATTAAGA | AATCTTCTAA | 180 |
| CCAACTGATT | GCTGGAATAT | CGAGACCGTT | AGTAGGCTCA | TCCAGTAATA | GTACGTCTGG | 240 |
| TTCACCGAAT | AAACTTTGCG | CTAATAATAC | TTTAATTTTT | TGGTTGTTTT | CTAATTCAGC | 300 |
| CATTTTTTTA | TCGTGTAAAG | TTGGATCGAT | ACCTAAACCA | GATAAAAGGT | TAGCAGCATC | 360 |
| AGCTTCAGCA | TTCCAACCAT | TCATTTCTGC | AAATTCACCT | TCAAGTTCAG | CAGCACGGAT | 420 |
| ACCATCTTCA | TCACTGAAAT | CTGGCTTCAT | ATAGATTTCA | TCTTTTTCTT | TCATAACCTC | 480 |
| ATAAAGACGT | TCGTGACCTT | TAATTACAAC | ATCAAGCACG | CGTTCATCTT | CATAAGCATA | 540 |
| GTGGTCCTGT | TTTAAAACAG | CTAGACGTTC | ATTTTTCCCT | AATGAAACAT | GTCCTGTTTG | 600 |
| AGAATCTAAT | TCACCAGATA | ATATTTTTAA | GAATGTTGAT | TTACCTGCAC | CATTCGCACC | 660 |

| | ATCTCCAAAA | CGTAAACTCA | CATCAGTTAC | TTGTAACATG | CATTTTCTCC | TTTTTTTCAT | 780 |
|-----|------------|------------|------------|-------------|------------|------------|------|
| | TCGATATTCT | AACGGAAGAA | TTATATCATA | TTATCGTCAC | AGTTTCGACC | TCATATAAGT | 840 |
| 5 | TGTAATGATA | GAATGACTCA | CACATGTTAT | AATAATAAAG | AATACAAGAA | TCGAAGGAGA | 900 |
| | ATAACATGGC | ATTAGACAAA | GATATAGTAG | GTTCTATAGA | ATTCCTTGAA | GTAGTAGGGT | 960 |
| | TACAAGGTTC | AACTTACCTT | TTAAAAGGAC | CAAACGGTGA | AAACGTAAAG | TTAAACCAAT | 1020 |
| 10 | CAGAAATGAA | CGATGATGAT | GAATTAGAAG | TAGGTGAAGA | ATATAGTTTC | TTCATTTATC | 1080 |
| | CAAACCGTTC | AGGTGAATTA | TTTGCAACTC | AAAATATGCC | TGATATTACG | AAAGATAAAT | 1140 |
| 15 | ATGACTTTGC | TAAAGTACTT | AAAACGGATC | GCGATGGGGC | ACGTATAGAT | GTTGGATTAC | 1200 |
| | CCCGTGAAGT | GTTAGTACCA | TGGGAAGATT | TACCAAAAGT | GAAATCACTA | TGGCCACAAC | 1260 |
| | CTGGTGATTA | TTTGCTAGTT | ACATTACGAA | TTGACCGTGA | GAATCATATG | TATGGACGTT | 1320 |
| 20 | TAGCGAGTGA | ATCTGTTGTA | GAAAATATGT | TTACACCTGT | ACACGACGAT | AAAATTTAA | 1380 |
| | ACGAAGTCAT | TGAAGCCAAA | CCTTACCGCG | TATTACGAAT | TGGTAGCTTT | TTATTAAGCG | 1440 |
| | AATCAGGTTA | CAAAATTTTC | GTACATGAAT | CAGAACGTAA | AGCTGAACCA | AGATTAGGTG | 1500 |
| ?5 | AATCTGTTCA | AGTTAGAATT | ATCGGGCATA | ATGATAAAGG | TGAGTTAAAT | GGTTCATTTT | 1560 |
| | TACCACTTGC | ACATGAACGT | TTAGACGATG | ACGGCCAAGT | CATCTTTGAT | TTACTAGTTG | 1620 |
| | AATATGATGG | TGAATTACCA | TTCTGGGACA | AATCAAGCCC | TGAAGCGATT | AAAGAAGTAT | 1680 |
| 30 | TCAATATGAG | TAAAGGTTCA | TTCAAACGTG | CAATCGGTCA | CTTATATAAA | CAGAAGATTA | 1740 |
| | TTAATATAGA | AACAGGTAAA | ATCGCTTTAA | CTAAAAAAGG | TTGGAGTCGA | ATGGACTCAA | 1800 |
| ne. | AAGAATAATC | ATTTTTACAC | GTGTCGTAGG | ATGCGTGTTT | TTTTTATTCA | ATATTAAATC | 1860 |
| 35 | GGACAGATGA | AGTAGTTTTT | TAAACATTCC | TTTCAAAGTA | AAAAATTAAA | TAATTCAAAC | 1920 |
| | GAATĀGGCTG | GGaCATTAAG | TTCTTAGGCA | ATGTAAAAAA | GCTGATTTCT | ATTAATTATT | 1980 |
| 10 | TGATGGAAAT | CAGCTTTTTT | GATATGTATT | TTATAATGTA. | CAGCTCGTTG | AGCTGCTATT | 2040 |
| | TTCCTTATAT | TAAGTGCCAT | TAATACAAAA | CCTAGCTCTC | GTTTAACTTT | ATTTATTCCT | 2100 |
| | CGAACTGACA | TTCGAGTGAA | aCCCAAAATA | GCCTTCATAA | ATCCAAAAAC | AGGCTCTACA | 2160 |
| 15 | TAAATTTTTC | TATGACTATA | GATTTTTTTC | GTTTCTGGTT | CAGAAAGCTT | TTGaTTAATT | 2220 |
| | TGGGCTTTAA | TGTATTTCAA | AGTAAAATTA | CATGTTAATA | CGTAGTATTA | ATGGCGAGAC | 2280 |
| | TCCTGAGGGA | GCAGTGCCAG | TCGAAGACAG | GGGCCCCAAC | ACAGAAGCTG | ACATATAGTC | 2340 |
| 50 | AGCTTACAAC | AATGTGCCGG | TTGGGGTGGC | TGAGACGGCA | CCCTAGGAAG | GGACCCGTCA | 2400 |
| | TCAAAAATTC | TATTTATAGA | ATTTTACAGT | AATGTGACAG | ACGGGCAAAG | CGAAGCCATT | 2460 |

| | CITACIGCIG | TTTTTTTAGG | GATTTATGTC | CCAGCCATTT | TTGTATTCAT | ATTTAAATTT | 2580 |
|------------|------------|-------------|------------|------------|------------|------------|------|
| | CGATAATTTT | TCAGGAAGCA | TTTTAATTTT | ACTAATGAAG | CAATATTTTT | TAGATTAACA | 2640 |
| 5 | AAAATTAATA | TTTACATTTT | CTTAACAATT | TTTTATGTAA | CATTTACAGT | TTCTAAAAAT | 2700 |
| | GAGGTTAATA | ATTCAAGGTT | AAGATAAAGA | TGTAATCAAT | ACAAATACTA | TTTGTTGTTC | 2760 |
| | ATACAGGGAG | GATATTTCAA | TGAAAAAATG | GCAATTTGTT | GGTACTACAG | CTTTAGGTGC | 2820 |
| 10 | AACACTATTA | TTAGGTGCTT | GTGGTGGCGG | TAATGGTGGC | AGTGGTAATA | GTGATTTAAA | 2880 |
| | AGGGGAAGCT | AAAGGTGATG | GCTCATCAAC | AGTAGCACCA | ATTGTGGAGA | AATTAAATGA | 2940 |
| 15 | AAAATGGGCT | CAAGATCACT | CGGATGCTAA | AATCTCAGCA | GGACAAGCTG | GTACAGGTGC | 3000 |
| | TGGTTTCCAA | AAATTCATTG | CAGGAGATAT | CGACTTCGCT | GATGCTTCTA | GACCAATTAA | 3060 |
| | AGATGAAGAG | AAGCAAAAAT | TACAAGATAA | GAATATCAAA | TACAAAGAAT | TCAAAATTGC | 3120 |
| 20 | GCAAGATGGT | GTAACGGTTG | CTGTAAATAA | AGAAAATGAT | TTTGTAGATG | AATTAGACAA | 3180 |
| | ACAGCAATTA | AAAGCAATTT | ATTCTGGAAA | AGCTAAAACA | TGGAAAGATG | TTAATAGTAA | 3240 |
| | ATGGCCAGAT | AAAAAATAA | ATGCTGTATC | ACCAAACTCA | AGTCATGGTA | CTTATGACTT | 3300 |
| 25 | CTTTGAAAAT | GAAGTAATGA | ataaagaaga | TATTAAAGCA | GAAAAAAATG | CTGATACAAA | 3360 |
| | TGCTATCGTT | TCTTCTGTAA | CGAAAAACAA | AGAGGGAATC | GGATACTTTG | GATATAACTT | 3420 |
| | CTACGTACAA | AATAAAGATA | aattaaaaga | AGTTAAAATC | AAAGATGAAA | ATGGTAAAGC | 3480 |
| 30 | AACAGAGCCT | ACGAAAAAA | CAATTCAAGA | TAACTCTTAT | GCATTAAGTA | GACCATTATT | 3540 |
| | CATTTATGTA | AATGAAAAAG | CATTGAAAGA | TAATAAAGTA | ATGTCAGAAT | TTATCAAATT | 3600 |
| 25 | CGTCTTAGAA | GATAAAGGTA | AAGCAGCTGA | AGAAGCTGGA | TATGTAGCAG | CACCAGAGAA | 3660 |
| 35 | AACATACAAA | TCACAATTAG | ATGATTTAAA | AGCATTTATT | GATAAAAATC | AAAAATCAGA | 3720 |
| | CGACAAGAAA | TCTGATGATA | AAAAGTCTGA | AGACAAAAA | TAATAAGACG | CAATTTCAAA | 3780 |
| 10 | TGTGTCTTGA | AACATGATTT | TGATGGTGAA | TCATTATTTA | GAGTACAAAG | CTTGATTTAT | 3840 |
| | CGAGACGCTG | ATTTTGACAT | TCAGTTAGTC | TACAAGCTTA | TCAACTTAAA | ATAGTGGTTC | 3900 |
| | ATCATTATTT | TACAAATCTA | ATTATTTTGG | GAGTAATAGA | AAGAGGTTTG | ATTATGACTT | 3960 |
| 1 5 | CATCTACTAA | TGTTAAAGCT | TTAATCGAAA | AAAATAATAA | TAAAAAAGGA | AAGCATAATG | 4020 |
| | ACAAAATTAT | ACCAGTTATT | TTAGCCGCAA | TTTCAGCGAT | TTCCATTTTA | ACAACACTAG | 4080 |
| ٠ | GTATATTAAT | CACATTGCTT | TTAGAAACCA | TCACTTTTTT | CACCAGAATT | CCAATAACTG | 4140 |
| 5 <i>0</i> | AATTTCTATT | TTCTACTACT | TGGAATCCTA | CCGGTTCAGA | CCCTAAGTTT | GGTATCTGGG | 4200 |
| | C100010110 | 1.0001.0000 | | | | | |

| | AACCGATATT | AGAAATTTTA | GCAGGAATAC | CAACAATTGT | GTTTGGTTTC | TTTGCATTAA | 4380 |
|----|------------|------------|------------|------------|------------|------------|------|
| | CCTTTGTTAC | ACCAGTATTA | AGATCTTTCA | TACCAGGTCT | TGGAGAGTTT | AATGCTATAA | 4440 |
| 5 | GTCCCGGCTT | AGTTGTCGGT | ATTATGATTG | TCCCTCTCAT | CACAAGTTTG | AGTGAGGaTG | 4500 |
| | CAATGGCATC | TGTACCAAAT | AAAATTCGAG | AAGGTGCCTA | TGGACTTGGA | GCAACTAAAT | 4560 |
| | TAGAAGTAGC | AACTAAAGTC | GTACTTCCCG | CAGCAACATC | AGGTATTGTA | GCTTCAATCG | 4620 |
| 10 | TTCTCGCGAT | TTCAAGAGCA | ATTGGAGAAA | CGATGATTGT | ATCATTAGCG | GCAGGTAGTT | 4680 |
| | CGCCAACAGC | TTCATTAAGT | TTAACAAGTT | CGATTCAAAC | AATGACTGGA | TATATTGTTG | 4740 |
| 15 | AGATAGCGAC | AGGTGATGCA | ACATTTGGAT | CAAATATTTA | TTACAGTATT | TATGCTGTAG | 4800 |
| | GGTTCACACT | ATTTATCTTT | ACCTTAATCA | TGAATTTACT | TTCTCAGTGG | ATTTCTAAGC | 4860 |
| | GTTTTAGGGA | GGAGTATTAA | TATGGAAACG | ACAGATAATA | ATAGACAATC | ACTCGTCGAT | 4920 |
| 20 | CAACAACTTG | TCCAAAAACA | TTTATCATCC | AGAĄCGGTTA | AAAATAAAGT | GTTCAAACTC | 4980 |
| | ATATTTTTAG | CATGTACATT | ATTAGGACTT | GTCGTACTTA | TTGCGTTGTT | AACTCAAACA | 5040 |
| | TTGATTAAAG | GGGTAAGTCA | TTTAAATTTA | CAGTTTTTCA | CTAATTTTTC | TTCTTCAACA | 5100 |
| 25 | CCATCTATGG | CTGGCGTTAA | AGGCGCGTTA | ATCGGTTCAC | TTTGGTTAAT | GTTAAGTATC | 5160 |
| | ATTCCATTAT | CAATCATCCT | AGGAATAGGT | ACAGCTATAT | ACTTAGAAGA | ATATGCGAAA | 5220 |
| | AACAACAAAT | TTACTCAGTT | TGTTAAAATC | AGTATTTCCA | ATTTAGCTGG | TGTACCATCA | 5280 |
| 30 | GTTGTATTTG | GGTTATTAGG | TTATACTTTG | TTCGTTGGTG | GTGCAGGGAT | TGAAGCCTTG | 5340 |
| | AAAATGGGTA | ACAGTATATT | GGCAGCAGCG | CTAACAATGA | CCTTACTGAT | ATTACCAATT | 5400 |
| 25 | ATTATTGTTT | CAAGTCAGGA | AGCAATTAGA | GCTGTACCTA | ACTCAGTACG | CGAACTTCTT | 5460 |
| 35 | ACGGCTTAGG | TGCTAATAAA | TGGCAAACGA | TAAGACGTGT | TGTCTTACCA | GCAGCGTTAC | 5520 |
| | CTGGTATTTT | AACTGGATTC | ATTTTGTCTC | TTTCAAGAGC | ACTGGGAGAA | ACAGCGCCAC | 5580 |
| 40 | TTGTGCTAAT | CGGTATACCG | ACTATATTAT | TGGCAACACC | TAGAAGTATA | TTGGATCAAT | 5640 |
| | TTTCAGCATT | ACCTATCCAA | ATATTTACTT | GGGCGAAAAT | GCCTCAAGAA | GAATTCCAGA | 5700 |
| | ATGTTGCATC | GGCAGGCATT | ATCGTTTTAC | TAGTTATCTT | AATCTTAATG | AATGGCGTTG | 5760 |
| 45 | CGATTATTTT | ACGTAACAAA | TTTAGTAAAA | AATTCTAATT | TAAACAATCA | ATCTCATTTA | 5820 |
| | TCTATTAAAA | AGGGAGTTTT | AAATATGGCG | CAAACACTTG | CACAAACTAA | ACAAATATCT | 5880 |
| | CAAAGTCATA | CGTTTGATGT | CTCACAAAGT | CATCATAAAA | CACCAGATGA | TACAAACTCA | 5940 |
| 50 | CATTCTGTTA | TATATTCAAC | ACAAAATTTA | GACTTATGGT | ATGGCGAAAA | TCATGCATTA | 6000 |
| | CAAAATATTA | АТТТАСАТАТ | TTATGAAAAC | САВАТТАСТС | CCATTATAGG | TCCATCTGGT | 6060 |

| | AAAACAGCTG | GTAAAATATT | ATATCGAGAT | CAAGACATTT | TTGATCAAAA | ATATTCTAAA | 6180 |
|-----------------|------------|------------|------------|------------|--------------------|------------|-------|
| | GAACAATTAC | GTACAAATGT | GGGCATGGTC | TTTCAACAAC | CTAATCCATT | TCCAAAATCA | 6240 |
| 5 | ATATACGATA | ATATTACTTA | CGGTCCAAAG | ATTCACGGTA | TTAAAAATAA | AAAAGTTCTT | 6300 |
| | GATGAAATCG | TTGAGAAATC | ATTACGTGGC | GCTGCAATTT | GGGATGAATT | AAAGGATAGG | 6360 |
| | TEGCACACAA | ATGCATATAG | TTTATCCGGT | GGGCAACAAC | AACGTGTTTG | TATCGCGCGT | 6420 |
| 10 | TGTTTAGCAA | TTGAACCTGA | AGTCATTTTA | ATGGATGAAC | CGACATCAGC | ATTAGATCCA | 6480 |
| | ATCTCAACAT | TAAGAGTAGA | AGAGTTGGTT | CAAGAACTAA | AAGAAAAGTA | TACAATTATT | 6540 |
| 15 | ATGGTLACAC | ATAATATGCA | ACAAGCAGCT | CGTGTATCAG | ATAAAACTGC | ATTTTTCTTA | 6600 |
| | AATGGTTATG | TCAATGAATA | TGATGATACT | GATAAAATTT | TCTCTAACCC | ATCAAACAAG | 6660 |
| | AAAACAGAAG | ATTATATTTC | AGGAAGGTTT | GGTTGATATA | TAATGGCAAT | AATTAGACAA | 6720 |
| 20 ⁴ | CGATATCAGG | AGCAACTTGA | TGATTTAATA | AAAGAATTAC | GTCGGTTAGG | TGCaAATGTC | 6780 |
| | TATGTGAGTA | TTGaAAATGG | TATAAAAtCA | TTAAGTATTG | aCGATAGAGG | CTTTGCACGA | 6840 |
| | CAAACAGTTA | AAAACGATAA | ACATATCAAT | CAATTAAATT | ATGATATTAA | TGAGCGAGTT | 6900 |
| 25 | ATCATGTTAA | TTACAAAGCA | ACAGCCCATT | GCGAGTGATT | TGCGTATGAT | GATTTCTTCA | 6960 |
| | TTAAAAATCG | CCTCCGATTT | AGAAAGAATA | GGAGATAATG | CCTCGAGTAT | TGCCAATATT | 7020 |
| | CGATTGCGTA | CAAAGATTAC | AGATGATTAT | GTGTTAACCC | GTTTAAAGAC | AATGGGTAAA | 7080 |
| 30 | TTAGCTATGT | TAATGTTAAA | GGACTTAGAT | CAAGCATTTA | AAAAGAAAGA | TACCGTATTA | 7140 |
| | ATAAGAGAAA | TAATTGAGCG | TGATGAAGAT | ATCGATGACT | TATATAGTCA | TATTATTAAC | 7200 |
| | GCAACGTATC | TTATTGATAA | CGtCCATTTG | TCGCTGCACA | AGCTCATTTA | GCAGCAAGAC | 7260 |
| 35 | ATTTAGAACG | TATTGGTGAT | CATATTATTA | ACATCGCTGA | AAGTGTTTAT | TTTTATTTAA | 7320 |
| | CAGOTACACA | TTACGAACAA | TAACTTAAAG | TTATTACTAT | AAAATCCCTT | ACGATAAATA | .7380 |
| 40 | TATATTTCTA | TTATTCATAA | ACCCTCAAAA | AAACCAAGAT | TCTCACAATT | AGTAATGTGA | 7440 |
| | AAATCTTGGT | TTATATTGTT | СТАСТАТААА | TTGTCTCGCA | TCTTAGTTAT | TTGCTTGCTC | 7500 |
| | AATTTCATCT | GTTAATTTTT | CAACTTCATC | GACTAAATCA | GAAATATATT | GAATTGTAGA | 7560 |
| 45 | TTTAAGTGGC | TGTTCTGTAG | TAATGTCTAC | ACCTGCAATG | TTTGCAAGTT | CGACAGGTGA | 7620 |
| | TACACTACCA | CCTTTTTTCA | ATGTTTCTAA | CCAAGCATCA | ACAGCTGGTT | GGCCTTCATT | 7680 |
| | TTTAATCTTT | TGAGAAACGA | CAGTTCCGAT | TGTTAAGCCA | GCAGAATACG | TATACGAATA | 7740 |
| 50 | TAATCCCATA | TAGTAATGAG | GTTGACGCAT | CCATGTTAAT | TCAGCACCCT | CAGTCATGTC | 7800 |
| | тастссатет | CCABABATT | СТТАТАВАС | ATTACCATT | מ דידיר מ יידידי מ | ATGTTCCCCC | 7860 |

(2) INFORMATION FOR SEQ ID NO: 162:

| 5 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 798 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
|------------|---|-----|
| 10 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 162: | |
| | TTTTTCTTT TCTTCATTTG AAAATTGATC ATTCAGCAAT ATAAGCGTAT TTGTTAATGA | 60 |
| 15 | TTTAGGTGTT CCAATTTCAT AATCCCACCA ATTTAAGTTG GTATTCTTGC CAGTTGTTTT | 120 |
| | AGTAAAATTC TCACTTAATT CTTTTACTTT TTTATCTGGT TCTTTTCCAT ATGCATTTTT | 180 |
| | ATGCAGCCAC TCAAGGGCAT CTTTCACTTT CTTCTTATTT TCGTCAGTAT TTAAAGTGGT | 240 |
| 20 | TTTAGGATTC CTCATCGCTT CTGCGATTTT CTCAATATTA CGATAGGTAC GAGTCATATG | 300 |
| | AGAAGAATTA GTTTCAAGGG TTTCCGCTCC TGACCACAAG TATTTCCTAC CACTTTCAGT | 360 |
| | TTTCATTTCC TTGAGTAAAT TCGTCGCCTC TTTCTCTGTA GCATCAAACT TCTTCTTCAT | 420 |
| 25 | ATCTGGATTA TTCTCATCAT ACTTATCATA ACCATAGTTA ACGTCCAGCC ATGTGTTCCT | 480 |
| | CAATTTTCA TAATCTGGCG TTTGAACATT CGTATCAGCC ACAGCGATTT GATGTTTATC | 540 |
| | AACACTTCTG AATTCACCAC CATTCAAAGT AATCACACCA GCCATTAATA ACGTAATGGT | 600 |
| 3 0 | GGATAATTIT TGCCATTTCT TTATTCTATA TGTCATTGAC ATGTCTCCTT TTTGTGTTGC | 660 |
| | GCGTGCGCAA TGAATATTAT GATTAAATAA TGATTCAAATT TTTCAAAATT CGTTAACGTA | 720 |
| | TACAAATGAC TGTCTACTGT CAAACAATCC ACAAAGAATG TTGATGECAT ATAAACAATC | 780 |
| 35 | GATCACCCAA ATTTTCCG | 798 |
| | (2) INFORMATION FOR SEQ ID NO: 163: | |
| 40 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 5132 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 45 | | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 163: | |
| | TACAGGTTTT ACTATAATGG ATGGTATTTT GGCTAAACGA CATTGGTTTA GTCTTCTTTT | 60 |
| 50 | TTTMACTICC TAMATITACA ATGGTATAAA TAATAATGCT ATATTTAGAA TGATGAGTAT | 120 |
| | ACTTACTGAA ACTAAATTAA AAGTGTCTGG TTCTTTACTA AAGATAGCTG CTATCCTTGC | 180 |

| | AATACAAGTT | CCAATGAGCG | CAATTAAAAG | TACTAACCCA | ACGATGAAAC | TCTGTTTGTC | 300 |
|----|------------|------------|------------|-------------|--------------|--------------|------|
| | ACTTAACTCA | AAGAAACTAT | AGATAGGATA | TTTTTTAATA | ATCAAGCCAC | CTAAAATCAT | 360 |
| 5 | CCATAAAAAT | ACGATAATTC | CATAAGTCAC | ATTTATAACA | TACGTTATTT | TTTGGTCACC | 420 |
| | AAATCGGACT | AATGTATTTC | GTAGAATCAG | CATACCAATG | ACAACACCTA | AAATAACGAT | 480 |
| - | ACTAGCTATA | TAAAGTAAAA | ATGCAATTGT | CACATCAAAT | GTACCCAAAT | CTAAAAACCT | 540 |
| 10 | AGGAATTAYA | AyGACTGCTA | AAATAAAAGC | GAAGYACAAA | GTAATATAKT | TATACAAACC | 600 |
| | GGTAGTAAGA | CTTATCTCAG | GTGATAATTG | ATCAGCCATT | GACTTAATCG | GTGTATTAAT | 660 |
| 15 | AATTGAACTT | GTATCTTCGT | TATTTTTTC | AGCCATAGTT | AAATGATCTT | CGAGCTCTTC | 720 |
| ,, | CAATAACTCT | TCTACTTCTG | CTTCAGTCTT | ACCTCTAAAT | AACAATTCAA | CACGTAATTT | 780 |
| | TTCTAAAAAA | TCTTGAGATT | GTTTACTTAA | CATCGTTTTC | CCCTCCAAAC | AAGTTAATCA | 840 |
| 20 | TCCCTTTATT | CAAAACTTGC | CATTTCGATT | TAAATACTTT | TAGTTCCTTT | AAACCTGAAT | 900 |
| | CGGTAATCGT | ATAGTATTTC | CGCCTCGGGC | CGCCATTACT | AGATTTTTT | ATTGTCGTAT | 960 |
| | CAACGTATCC | TITTTTGTTT | AAACGCATTA | AAACTGGATA | AATACTACCC | TCACTTATCT | 1020 |
| 25 | CTGGAAACTC | TTGATTCTTA | AGTTTCGTCA | TAATTTCATA | TCCATACGTT | TCGCCTTGGG | 1080 |
| | CAATGAGACC | TAATATCGCC | CCATCTAAGA | GACCTTTCAT | AATCTGATCT | GACACTGACA | 1140 |
| | TTTTAATCAC | CTACTATCTT | ACATAATAAG | ATAGTACATT | GAGAACTTTT | CGTCAACTAT | 1200 |
| 30 | CTTTTATTGT | AAGGTAGTTG | TTGTACACAT | TCCTTAAATG | ACTAACAACT | TTGTTAATAG | 1260 |
| | GGTAATACTT | ACGGAAGTAT | ATTTTATTTA | TGGGGGAGGA | ATTAATAATG | ACTACAAAAA | 1320 |
| | CAGTATTTGA | TGTCATTGAT | ATGGGGTTAG | GATATTTAGT | AAATGTGTAT | GATGCTTGGA | 1380 |
| 35 | AAGTTGAAAA | GGTACTTGAT | GATTATCATA | AGCCTTTTTC | TAATACCATT | CATTGGCAAT | 1440 |
| | TTGGECATGT | ATTAACAATT | TTTGAATCGG | CCTTAGCTGT | TGCTGGTAAA | GAGAATATTG | 1500 |
| 40 | ATTTAAATAT | CTATAGACCT | TTATTCGGAA | ATGGTTCGTC | TCCAGATGAA | TGGAAGGATG | 1560 |
| | AAGTACCGAG | TATTGAAAGG | ATTTTAGAAG | -GTCTCCAAAC | -TTTACCTGAA- | -CGTGCACGAA- | 1620 |
| | ATCTAACTGA | AGATGATTTA | GCAATTGAAT | TGAAACAGCC | AATTGTCGGT | TGTAATAACT | 1680 |
| 45 | TAGAAGAGTT | ATTAGTATTA | AATGCCATTC | ACATCCCACT | TCATGCTGGT | AAAATTGAAG | 1740 |
| | AGATGTCTCG | TATATTAAAA | TAAAATTTAA | AAATATGTGC | TTATTAACCG | TTAACAACAC | 1800 |
| | GTTAACGGgT | TTTTTATTTG | TTTAAAAGGT | CACTTTTTTG | AATTTAATAA | ACACCATCTA | 1860 |
| 50 | TACCAGTTCT | TCACCGATTC | TCGAAAAATA | ATTATATTAA | TGATTTCGTT | AATTTAATTT | 1920 |
| | TATATTTAAT | TATTACTGTA | CATCTTTTGT | AGTTAGCTTT | ATTCTTAAAT | TGAAATATGT | 1980 |

| | TACTCCCTAT | CGTTGTAGGT | CTCCTTATTT | GGGCACTTAC | ACCTTTTAAA | CCGGATGCTG | 2100 |
|------------|-----------------|-------------|----------------|------------|--|------------|------|
| | TGGATCCAAC | AGCATGGTAT | ATGTTCGCAA | TATTCGTCGC | GACAATCATT | GCTTGTATTA | 2160 |
| 5 | CACAACCGAT | GCCAATTGGG | GCCGTCTCTA | TAATTGGATT | TACAATCATG | GTACTCGTTG | 2220 |
| | GCATTGTTGA | CATGAAAACG | GCTGTCGCTG | GTTTTGGTAA | TAATAGCATT | TGGTTAATTG | 2280 |
| _ | CTATGGCATT | TTTCATTTCG | AGAGGATTTG | TGAAAACAGG | TCTTGGTAGA | CGTATCGCAC | 2340 |
| 10 | TTCATTTCGT | CAAATTATTT | GGTAAAAAAA | CATTAGGATT | AGCATATTCT | ATCGTCGGTG | 2400 |
| | TAGATTTAAT | TCTAGCGCCT | GCTACACCAA | GTAATACCGC | GCGTGCTGGT | GGAATCATGT | 2460 |
| 15 | TCCCAATTAT | CAAATCACTT | TCTGAATCAT | TTGGTTCGAA | ACCGAAAGAC | GGATCAGCAC | 2520 |
| | GCAAAATGGG | TGCATTTCTT | GTTTTCACAG | AATTCCAAGG | TAATTTAATT | ACTGCGGCTA | 2580 |
| | TGTTTTTAAC | TGCAATGGCC | GGTAACCCCC | TTGCACAAAA | TTTAGCATCT | AGCACATCTA | 2640 |
| 20 | ATGTTCACAT | TACATGGATG | AATTGGTTTC | TAGCTGCTTT | AGTTCCTGGA | CTTGTTTCCT | 2700 |
| | TAATTGTTGT | ACCTTTTATT | ATTTATAAAA | TTTATCCACC | AACTGTTAAA | GAAACACCAA | 2760 |
| | ATGCTAAGAG | TTGGGCTGAA | AATGAATTAG | CGACTATGGG | TAAAATCGCT | TTAGCTGAAA | 2820 |
| ?5 | AATTTATGAT | TGGTATTTTT | GTCGTTGCGT | TAACACTATG | GATTGTCGGA | AGTTTCATTC | 2880 |
| | ATATTGATGC | AACTTTAACG | GCCTTTATTG | CGCTAgcATT | gttattattg | ACAGGCGTCT | 2940 |
| | TAACATGGCA | AGACATTTTA | AACGAAACAG | GTGCTTGGAA | CACATTAGTA | TGGTTCTCAG | 3000 |
| 30 | TATTAGTGTT | AATGGCCGAC | CAATTAAACA | AGCTTGGATT | TATTCCTTGG | TTAAGTAAAT | 3060 |
| | CCATTGCTAC | AAGTCTTGGT | GGCTTAAGCT | GGCCTATAGT | CCTGGTCATT | TTAATATTGT | 3120 |
| 25 | TCTACTTCTA | TTCACATTAC | TTATTTGCAA | GTTCTACAGC | ACATATCAGT | GCGATGTATG | 3180 |
| 35 | CAGCATTACT | AGgCGTTGCC | ATCGCAGCCG | GTGCACCACC | ATTATTCAGT | GCATTAATGT | 3240 |
| | TAGGTTTCTT | CGGTAACCTA | TTAGCTTCAA | CAACACACTA | TAGTAGTGGT | CCAGCGCCGA | 3300 |
| 40 | TTCTATTCTC | TTCAGGTTAC | GTGACTCAAA | AACGTTGGTG | GACAATGAAC | TTAATATTAG | 3360 |
| | GTTTCGTCTA | CTTTATTATC | TGGATTGGTT | TAGGATCACT | TTGGATGAAA | GTAATTGGTA | 3420 |
| | TATTTTAAAA | TATTTAAATT | AGCGCTCGAA | TCTCATTGAT | TTGGGCGCTT | TTTAATTTGT | 3480 |
| 45 | ATTTAAAATC | AACCTTTGCT | AAATCAAGAC | TCCCTTTTTA | AAATACGTTT | ATCCTTTAAA | 3540 |
| | TCATTGCGTG | CTTCACTGAA | AATTTGTATA | AAGATTTAAG | TCATTACGTA | ACATCACATA | 3600 |
| | AAATACATTT | CTATACTATT | CCGCTTCATT | GATTAACATT | ACGTATGCCC | TCATAAATCA | 3660 |
| 5 <i>0</i> | TCATACAAAA | AACACCTTCG | TTTAAATTCA | TTTTAATTGC | GAATTCAACG | AAAGTGCCTT | 3720 |
| | 1 mmm (1 m 1 mm | # 1 mcmmc 1 | 3 3 0000 D 000 | ********** | TOTA (************************************ | CATACCTCAC | 3780 |

| | TTATAGGGTT | TTTGCGACCG | GATGTTTCTT | CAATTTAATG | TATTGAGAAA | GACTATATAA | 3900 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | CACAATACCT | GTCCAAATAA | ATATAAACGT | AATTAATTGA | TCTATACTAA | AAGGCTCTTT | 3960 |
| 5 | GAAAACAAAT | ATGCCGAGTA | CAAACATTAT | TGTTGGTCCA | ACGTATTGAA | TAAATCCTAT | 4020 |
| | TAGCGAAAGT | GGAATACGTT | TTGCCCCGGC | TGAGAATAGG | ATTAGTGGTA | TTGCCGTAAT | 4,080 |
| _ | AGCÁCCAGAA | AATAACAACC | AAAATGATGA | CATGTTCAAT | CCAAATGACA | TCTGATGTTG | 4140 |
| 10 | CTGCCATAAA | TAAATAACGT | ATATTAGTCC | AGCAGGTGCG | GTAACAATAC | ATTCAATCGT | 4200 |
| | AATACTGCTG | ATGGCATCAA | TATGTACTAC | TTTTTTCAAT | AATCCGTATG | TACCAAAGGA | 4260 |
| 15 | TAACGCTAAT | ATAATAGAGA | CGATTGGGAA | TTCTCCAATC | TTGAGCGTCA | TATATAATAC | 4320 |
| | ACCGATGAAT | GCGAATAAAA | TGGCTAGCCA | TTCAAATTTA | TTGAATCTTT | CTTTTAAAAA | 4380 |
| | GATAAGTGCG | AGCAAAATGC | TAACAAGTGG | ATTTATATAA | TAACCTAAAC | TTGTTTGTAG | 4440 |
| 20 | GACGTGACCG | TTCGTTACAG | CCCAAATAAA | TGTACCCCAA | TTTAATGTAA | TGACATAGCC | 4500 |
| | TGCTACGACA | ATCGCTAATA | GCTGAATGGG | CTTGCCTAAC | AATTGATTCA | TATCTCGTTG | 4560 |
| | AAATGCATTG | CGTTGTTTTT | GTCCAACCGC | GAGTATGAAA | ATCATGAATA | TTGCTGAAAA | 4620 |
| 25 | TATAATACGA | AAGGCTAAAA | TTTCAAATGC | GCCTATTGCA | TCAACGAACT | GCCAATATAT | 4680 |
| | AGGTAGTATT | CCCCACAGAA | TGTATGCACT | GAGTGCTAAA | AATATGCCTT | TTTTATACTC | 4740 |
| | TGAATTCACC | TTCAAACCTC | CTTACTTTCC | TAATTTTTAA | TTTACTGCAT | ACGCTCACTT | 4800 |
| 30 | GGTTATGCTA | ATATAACGAT | TTTACTAATA | ATATTTCGAT | AAAGATATCA | TTTTGTTTAT | 4860 |
| | ATTTCCCACA | TTTATTCACC | AACCACTAAA | CAATATTAAT | TTTATAAATA | ATTCTGTACA | 4920 |
| | AATCAGGGTA | TATTGCCAGA | AAGACTACCA | TACAACATAA | AGGATGGATA | CAAATGACTT | 4980 |
| 35 | TACCTAAAAT | TGGAAAGCCT | GCAACACGCG | CGCTAAATTC | ACAAGGTATA | TACACATTAG | 5040 |
| | AAGÇÃGTATC | ACAATATACG | AAGTCATCTC | TAATGGAGAT | GCATGGCGTT | GGTCCTAAAG | 5100 |
| 40 | CTATATCAAT | ATTGGAACAA | GCTTTATTTC | AG | | | 5132 |
| | | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 164:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22243 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 164:

AAGTAAATTA TATTATGAAT TTGCCTGTCA ATTTCTTAAA GACATTCTTA CCGGAACTAA 60

55

45

| • | TAGAAGCAAT | TAATAATGCY | mAAGAAAAGA | CAGCTAATAA | TACCGGCTTA | AAATTAATAT | 180 |
|-----------|------------|------------|----------------|--------------|------------|------------|------|
| | TTGCAATTAA | TTATGGTGGC | AGAGCAGAAC | TTGTTCATAG | TATTAAAAAT | ATGTTTGACG | 240 |
| 5 | AGCTTCATCA | ACAAGGTTTA | AATAGTGATA | TCATAGATGA | AACATATATA | AACAATCATT | 300 |
| | TAATGACAAA | AGACTATCCT | GATCCAGAGT | TGTTAATTCG | TACTTCAGGA | GAACAAAGAA | 360 |
| 0 | TAAGTAATTT | CTTGATTTGG | CAAGTTTCGT | ATAGTGAATT | TATCTTTAAT | CAAAAATTAT | 420 |
| | GGCCTGACTT | TGACGAAGAT | GAATTAATTA | AATGTATAAA | AATTTATCAG | TCACGTCAAA | 480 |
| | GACGCTTTGG | CGGATTGAGT | GAGGAGTAGT | ATAGTATGAA | AGTTAGAACG | CTGACAGCTA | 540 |
| 5 | TTATTGCCTT | AATCGTATTC | TTGCCTATCT | TGTTAAAAGG | CGGCCTTGTG | TTAATGATAT | 600 |
| | TTGCTAATAT | ATTAGCATTG | ATTGCATTAA | AAGAATTGTT | GAATATGAAT | ATGATTAAAT | 660 |
| | TIGITTCAGT | TCCTGGTTTA | ATTAGTGCAG | TTGGTCTTAT | CATCATTATG | TTGCCACAAC | 720 |
| 20 | ATGCAGGGCC | ATGGGTACAA | GTAATTCAAT | TAAAAAGTTT | AATTGCAATG | AGCTTTATTG | 780 |
| | TATTAAGTTA | TACTGTCTTA | TCTAAAAACA | GATTTAGTTT | TATGGATGCT | GCATTTTGCT | 840 |
| | TAATGTCTGT | GGCTTATGTA | GGCATTGGTT | TTATGTTCTT | TTATGAAACG | AGATCAGAAG | 900 |
| ?5 | GATTACATTA | CATATTATAT | GCCTTTTTAA | TTGTTTGGCT | TACAGATACA | GGGGCTTACT | 960 |
| | TGTTTGGTAA | AATGATGGGT | AAACATAAGC | TTTGGCCAGT | AATAAGTCCG | AATAAAACAA | 1020 |
| | TCGAAGGATT | CATAGGTGGC | TTGTTCTGTA | GTTTGATAGT | ACCACTTGCA | ATGTTATATT | 1080 |
| 30 | TTGTAGATTT | CAATATGAAT | GTATGGATAT | TACTTGGAGT | GACATTGATT | TTAAGTTTAT | 1140 |
| | TTGGTCAATT | AGGTGATTTA | GTGGAATCAG | GATTTAAGCG | TCATTTCGGC | GTTAAAGACT | 1200 |
| ne | CAGGTCGAAT | ACTACCTGGA | CACGGTGGTA | TTTTAGACCG | ATTTGACAGC | TTTATGTTTG | 1260 |
| 35 | TGTTACCATT | ATTAAATATT | TTATTAATAC | AATCTTAATG | CTGAGAACAA | ATCAATAAAC | 1320 |
| | GTAĄĄGAGGA | GTTGCTGAGA | TAATTTAATG | AATCTCAGAA | CTCCTTTTGA | AAATTATACG | 1380 |
| 10 | CAATATTAAC | TTTGAAAATT | ATACGCAATA | . TTAACTTTGA | AAATTAGACG | TTATATTTTG | 1440 |
| | TGATTTGTCA | GTATCATATT | ATAATGACTT | ATGTTACGTA | TACAGCAATC | ATTTTTAAAA | 1500 |
| | TAAAAGAAAT | TTATAAACAA | TCGAGGTGTA | GCGAGTGAGC | TATTTAGTTA | CAATAATTGC | 1560 |
| 45 | ATTTATTATT | GTTTTTGGTG | TACTAGTAAC | TGTTCATGAA | TATGGCCATA | TGTTTTTTGC | 1620 |
| | GAAAAGAGCA | GGCATTATGT | GTCCAGAATT | TGCGATCGGT | ATGGGGCCAA | AAATTTTTAG | 1680 |
| | TTTTAGAAAA | AATGAAACAC | TTTACACTAT | TAGGTTATTG | CCTGTTGGTG | GATATGTTCG | 1740 |
| 50 | TATGGCAGGA | GATGGCTTAG | AAGAGCCACC | AGTCGAGCCC | GGTATGAACG | TTAAAATTAA | 1800 |
| | A | CAAAATCAAA | TN NC NC NT NT | САТАТТАСАТ | CATCATCATA | ACTITONACA | 1860 |

| | CACTGCTTAT | GATAATGAAA | GACATCATTT | TAAAATTGCT | AGAAAGTCTT | TCTTTGTTGA | 1980 |
|----|------------|------------|------------|------------|------------|------------|------|
| | AAATGGTAGC | TTAGTTCAAA | TTGCTCCGAG | AGACAGACAA | TTTGCACATA | AAAAGCCATG | 2040 |
| 5 | GCCGAAATTT | TTAACATTAT | TTGCGGGACC | GTTATTTAAC | TTTATATTAG | CTTTAGTCCT | 2100 |
| | ATTTATTGGT | CTTGCATATT | ATCAAGGCAC | GCCTACGTCT | ACTGTAGAAC | AAGTCGCAGA | 2160 |
| 10 | TAAGTATCCA | GCTCAACAAG | CAGGATTACA | AAAAGGTGAT | AAGATCGTCC | AAATTGGCAA | 2220 |
| 10 | ATATAAAATA | TCTGAATTTG | ATGATGTTGA | TAAGGCGTTA | GATAAAGTTA | AAGATAATAA | 2280 |
| | GACGACTGTT | AAATTTGAAC | GTGATGGTAA | AACAAAGTCA | GTTGAATTAA | CACCTAAAAA | 2340 |
| 15 | GACTGAAAAA | AAACTGACTA | AAGTAAGTTC | AGAGACGAAG | TATGTTCTCG | GATTCCAACC | 2400 |
| | AGCGAGTGAA | CATACACTTT | TTAAACCAAT | TGTATTCGGA | TTTAAAAGCT | TTTTAATCGG | 2460 |
| | TAGTACTTAT | ATTTTTACAG | CTGTAGTAGG | TATGTTGGCT | AGTATATTTA | CGGGCGGATT | 2520 |
| 20 | CTCATTTGAT | ATGTTAAATG | GTCCGGTTGG | TATTTATCAT | AACGTCGACT | CAGTTGTTAA | 2580 |
| | AGCGGGTATC | ATTAGCTTAA | TTGGTtnCAC | TGCGTTATTA | AGTGTAAACT | TAGGTATTAT | 2640 |
| | GAATTTAATT | CCTATTCCTG | CACTAGACGG | TGGTCGTATT | TTATTTGTTA | TATATGAAGC | 2700 |
| 25 | GATTTTCAGA | AAACCAGTTA | ATÄAAAAAGC | GGAAACAACG | ATTATTGCTA | TTGGTGCCAT | 2760 |
| | TTTCATGGTC | GTTATAATGA | TATTAGTAAC | GTGGAATGAT | ATTCGACGAT | ATTTCTTATA | 2820 |
| | ATTTAGGAGG | TTAAATAATT | ATGAAGCAAT | CCAAAGTTTT | TATACCAACG | ATGCGTGACG | 2880 |
| 30 | TGCCATCAGA | AGCAGAAGCA | CAAAGTCATC | GTTTATTATT | GAAATCGGGT | TTGATAAAAC | 2940 |
| | AAAGTACAAG | TGGGATTTAT | AGTTATTTAC | CGCTAGCAAC | ACGTGTGTTA | AATAATATTA | 3000 |
| 35 | CTGCAATTGT | GCGACAAGAA | ATGGAACGTA | TCGATTCTGT | TGAAATTTTA | ATGCCAGCGT | 3060 |
| | TACAACAAGC | TGAATTATGG | GAAGAATCAG | GACGTTGGGG | TGCATATGGC | CCAGAATTAA | 3120 |
| | TGCGTTTACA | AGATAGaCAT | GGAAGACAAT | TTgCATTAGG | TCCaACACAT | GAAGAATTAG | 3180 |
| 10 | TTACATCAAT | AGTAAGAAAT | GAATTGAAAT | CATACAAACA | ATTACCGATG | ACATTATTCC | 3240 |
| | AAATTCAATC | TAAATTCCGT | GATGAAAAGA | GACCACGTTT | TGGTTTAYTC | GTGGGCGTGA | 3300 |
| | ATTTATTATG | AAAGATGCAT | ATTCATTCCA | TGCTGACGAG | GCATCATTAG | ATCAAACGTA | 3360 |
| 15 | TCAAGATATG | TATCAAGCGT | ATAGCCGTAT | TTTTGAGAGA | GTTGGCATTA | ACGCAAGACC | 3420 |
| | AGTAGTTGCA | GATTCAGGTG | CTATAGGCGG | TAGCCATaCA | CATGAATTTA | TGGCATTAAG | 3480 |
| | TGCTATCGGT | GAGGATACAA | TCGTTTACAG | TAAAGAAAGT | GATTATGCTG | CTAACATCGA | 3540 |
| 50 | AAAAGCAGAA | GTCGTTTACG | ArcCAaATcA | Taagcatact | ACTGTGCAAC | CTTTAGAAAA | 3600 |
| | AATTGAAACA | ССАВАТСТТА | AGACTGCGCA | AGAATTCCCA | CACTTCTTAC | GTAGACCAGT | 3660 |

| | GCGTGGCCAT | CATGAAATTA | ATGACATTAA | ATTAAAATCT | TATTTCGGCA | CAGATAATAT | 3780 |
|------------|-------------|------------|------------|------------|------------|------------|------|
| | TGAATTAGCA | ACACAAGACG | AAATTGTTAA | TTTAGTTGGT | GCAAATCCTG | GTTCACTAGG | 3840 |
| 5 | TCCTGTAATT | GATAAAGAAA | TCAAAATTTA | TGCAGATAAT | TTTGTGCAAG | AATAAATTA | 3900 |
| | TTTAGTTGTC | GGTGCTAACG | AAGATGGTTA | TCACTTAATT | aatgtaaatg | TAGGTAGAGA | 3960 |
| | CTTCAACGTT | GATGAATATG | GCGATTTCCG | TTTTATTTTA | GAAGGCGAAA | AGTTAAGTGA | 4020 |
| 10 | TGGTTCAGGC | GTTGCACATT | TTGCTGAAGG | TATTGAAGTT | GGTCAAGTAT | TCAAATTGGG | 4080 |
| | TACTAAGTAT | TCAGAATCAA | TGAATGCTAC | ATTCTTAGAT | AACCAAGGAA | AAGCTCAATC | 4140 |
| 1.5 | TTTAATTATG | GGTTGTTACG | GAATTGGAAT | TTCTAGAACG | CTAAGTGCGA | TTGTTGAACA | 4200 |
| 15 | AAATCACGAT | GATAATGGAA | TTGTTTGGCC | TAAATCAGTT | ACTCCGTTTG | ATTTACATTT | 4260 |
| | AATTTCTATT | AATCCTAAGA | AAGATGATCA | ACGAGAACTA | GCAGATGCAC | TATATGCTGA | 4320 |
| 2 Q | ATTTAATACT | AAATTTGATG | TGTTGTACGA | TGATCGTCAG | GAACGTGCAG | GTGTTAAATT | 4380 |
| | TAATGATGCC | GATTTAATTG | GTTTACCACT | GCGAATTGTT | GTTGGTAAAC | GTGCATCGGA | 4440 |
| | AGGTATTGTA | GAAGTTAAAG | AACGTTTAAC | AGGTGATAGC | GAAGAAGTTC | ACATTGATGA | 4500 |
| 25 | CTTAATGACT | GTCATTACAA | ATAAATATGA | TAACTTAAAA | TAATTAAGAT | CGAATGAATT | 4560 |
| | ATAAGAGTAG | GAAAAAGCTG | AAAGAAATCT | GATGCTTATG | TCCTGCTCTT | ATTATTTTTG | 4620 |
| | ATATAATGAT | TATTCGATGA | AAAATGACTG | AAGACATAGT | ATAATTAAAG | ATAAATTTGT | 4680 |
| 30 | TTTAACAATA | TAATGATTAG | CCAAATATAA | AGCATTTAAT | TTTCTATCAT | TACTATGCTC | 4740 |
| | ACATAATCTA | AATATTGTTC | GAACACGTAA | AAGTAATTTC | TATTTAAGGT | GGTAATTGTC | 4800 |
| | TTGGCAATGA | CAGAGCAACA | AAATTTAAA | GTGCTTGCTG | ATCAAATTAA | AATTTCAAAT | 4860 |
| 35 | CAATTAGATG | CTGAAATTTT | AAATTCAGGT | GAACTGACAC | GTATAGATGT | TTCTAACAAA | 4920 |
| | AACAGAACAT | GGGAATTTCA | TATTACATTA | CCACAATTCT | TAGCTCATGA | AGATTATTTA | 4980 |
| | AATATTTATTA | ATGCAATAGA | GCAAGAGTTT | AAAGATATCG | CCAACGTTAC | ATGTCGTTTT | 5040 |
| 40 | ACGGTAACAA | ATGGCACGAA | TCAAGATGAA | CATGCAATTA | AATACTTTGG | GCACTGTATT | 5100 |
| | GACCAAACAG | CTTTATCTCC | AAAAGTTAAA | GGTCAATTGA | AACAGAAAAA | GCTTATTATG | 5160 |
| 45 | TCTGGAAAAG | TATTAAAAGT | AATGGTATCA | AATGACATTG | AACGTAATCA | TTTTGATAAG | 5220 |
| | GCATGTAATG | GAAGTCTTAT | CAAAGCGTTT | AGAAATTGTG | GTTTTGATAT | CGATAAAATC | 5280 |
| | ATATTCGAAA | CAAATGATAA | TGATCAAGAA | CAAAACTTAG | CTTCTTTAGA | AgCACaTATT | 5340 |
| 50 | CAAGAAGAAG | ACGAACAAAG | TGCACGATTG | GCAACAGAGA | AACTTGAAAA | AATGAAAGCT | 5400 |
| | | | | | | | EAR |

| | GCAATAGAGG | GTGTCATTTT | TGATATAAAC | TTAAAAGAAC | TTAAAAGTGG | TCGCCATATC | 5580 |
|----|------------|----------------|-------------|--------------|------------|-------------|-------|
| | GTAGAAATTA | AAGTGACTGA | CTATACGGAC | TCTTTAGTTT | TAAAAATGTT | TACTCGTAAA | 5640 |
| 5 | AACAAAGATG | ATTTAGAACA | TTTTAAAGCG | CTAAGTGTTG | GTAAATGGGT | TAGGGCTCAA | 5700 |
| | GGTCGTATTG | AAGAAGATAC | ATTTATTAGA | GATTTAGTTA | TGATGATGTC | TGATATTGAA | 5760 |
| | GAGATTAAAA | AAGCGACAAA | AAAAGATAAG | GCTGAAGAAA | AGCGTGTAGA | ATTCCACTTG | 5820 |
| 10 | CATACTGCAA | TGAGCCAAAT | GGATGGTATA | CCCAATATTG | GTGCGTATGT | TAAACAGGCA | 5880 |
| | GCAGACTGGG | GACATCCAGC | CATTGCGGTT | ACAGACCATA | ATGTTGTGCA | AGCATTTCCA | 5940 |
| 15 | GATGCTCACG | CAGCAGCGGA | AAAACATGGC | attaaaatga | TATACGGTAT | GGAAGGTATG | 6000 |
| | TTAGTTGATG | ATGGTGTTCC | GATTGCATAC | AAACCACAAG | ATGTCGTATT | AAAAGATGCT | 6060 |
| | ACTTATGTTG | TGTTCGACGT | TGAGACAACT | GGTTTATCAA | ATCAGTATGA | TAAAATCATC | 6120 |
| 20 | GAGCTTGCAG | CTGTGAAAGT | TCATAACGGT | GAAATCATCG | ATAAGTTTGA | AAGGTTTAGT | 6180 |
| | AATCCGCATG | AACGATTATC | GGAAACGATT | ATCAATTTGA | CGCATATTAC | TGATGATATG | 6240 |
| | TTAGTAGATG | CCCCTGAGAT | TGAAGAAGTA | CTTACAGAGT | TTAAAGAATG | GGTTGGCGAT | 6300 |
| 25 | GCGATATTCG | TAGCGCATAA | TGCTTCGTTT | GATATGGGCT | TCATCGATAC | GGGATATGAA | 6360 |
| | CGTCTTGGGT | TTGGACCATC | AACGAATGGT | GTTATCGATA | CTTTAGAATT | ATCTCGTACG | 6420 |
| | ATTAATACTG | AATATGGTAA | ACATGGTTTG | AATTTCTTGG | СТААААААТА | TGGCGTAGAA | 6480 |
| 30 | TTAACGCAAC | ATCACCGTGC | CATTTATGAT | ACAGAAGCAA | CAGCTTACAT | TTTCATAAAA | 6540 |
| | ATGGTTCAAC | AAATGAAAGA | ATTAGGCGTA | TTAAATCATA | ACGAAATCAA | CAAAAAACTC | 6600 |
| | AGTAATGAAG | ATGCATATAA | ACGTGCAAGA | CCTAGTCATG | TCACATTAAT | TGTACAAAAC | 6660 |
| 35 | CAACAAGGTC | TTAAAAATCT | TTAAAATTTA | GTAAGTGCAT | CATTGGTGAA | GTATTTCTAC | 6720 |
| | CGTAÉACCTC | GAATTCCACG | TTCATTGTTA | GATGAATATC | GTGAGGGATT | ATTGGTAGGT | 6780 |
| 40 | ACAGCGTGTG | ATGAAGGTGA | ATTATTTACG | GCAGTTATGC | AGAAGGACCA | GAGTCAAGTT | 6840 |
| | GAAAAAATTG | -CCAAATATTA | TGATTTTATT | -GAAATTCAAC- | CACCGCACT | TTATCAAGAT | -6900 |
| | TTAATTGATA | GAGAGCTTAT | TAGAGATACT | GAAACATTAC | ATGAAATTTA | TCAACGTTTA | 6960 |
| 45 | ATACATGCAG | GTGACACAGC | GGGTATACCT | GTTATTGCGA | CAGGAAATGC | ACACTATTTG | 7020 |
| | TTTGAACATG | ATGGTATCGC | ACGTAAAATT | TTAATAGCAT | CACAACCCGG | CAATCCACTT | 7080 |
| | AATCGCTCAA | CTTTACCGGA | AGCACATTTT | AGAACTACAG | ATGAAATGTT | AAACGAGTTT | 7140 |
| 50 | CATTTTTTAG | GTGAAGAAAA | AGCGCATGAA | ATTGTTGTGA | AAAATACAAA | CGAATTAGCA | 720 |
| | CATCCAATTC | A A COTTOTTOTT | TO COTATORA | CATCAATTAT | 303030000 | TATCCA ACCT | 200 |

| | CTGCCTCAAA | TCGTAATTGA | TCGATTAGAA | AAAGAATTAA | AAAGTATTAT | CGGTAATGGA | 7380 |
|----|------------|------------|------------|-------------|-------------|------------|------|
| | TTTGCGGTAA | TTTACTTAAT | TTCGCAACGT | TTAGTTAAAA | AATCATTAGA | TGATGGATAC | 7440 |
| 5 | TTAGTTGGTT | CCCGTGGTTC | AGTAGGTTCT | AGTTTTGTAG | CGACAATGAC | TGAGATTACT | 7500 |
| | GAAGTAAACC | CGTTACCGCC | ACACTATATT | TGTCCGAACT | GTAAAACGAG | TGAATTTTTC | 7560 |
| 10 | AATGATGGTT | CAGTAGGATC | AGGATTTGAT | TTACCTGATA | AGACGTGTGA | AACTTGTGGA | 7620 |
| 10 | GCGCCACTTA | TTAAAGAAGG | ACAAGATATT | CCGTTTGAAA | CATTTTTAGG | ATTTAAGGGA | 7680 |
| | GATAAAGTTC | CTGATATCGA | CTTAAACTTT | AGTGGTGAAT | ATCAACCGAA | TGCCCATAAC | 7740 |
| 15 | TACACAAAAG | TATTATTTGG | TGAGGATAAA | GTATTCCGTG | CAGGTACAAT | TGGTACTGTT | 7800 |
| | GCTGAAAAGA | CTGCTTTTGG | TTATGTTAAA | GGTTATTTGA | ATGATCAAGG | TATCCACAAA | 7860 |
| | AGAGGTGCTG | AAATAGATCG | ACTCGTTAAA | GGATGTACAG | GTGTTAAACG | TACAACTGGA | 7920 |
| 20 | CAGCATCCAG | GGGGTATTAT | TGTAGTACCT | GATTACATGG | ATATTTATGA | TTTTACGCCG | 7980 |
| | ÀTACAATATC | CTGCCGATGA | TCAAAATTCA | GCATGGATGA | CGACACATTT | TGATTTCCAT | 8040 |
| | TCTATTCATG | ATAATGTATT | AAAACTTGAT | ATACTTGGAC | ACGATGATCC | AACAATGATT | 8100 |
| 25 | CGTATGCTTC | AAGATTTATC | AGGAATTGAT | ССАААААСАА | TACCTGTAGA. | TGATAAAGAA | 8160 |
| | GTTATGCAGA | TATTTAGTAC | ACCTGAAAGT | TTGGGTGTTA | CTGAAGATGA | AATTITATGT | 8220 |
| | AAAACAGGTA | CATTTGGGGT | ACCAGAATTC | GGTACAGGAT | TCGTGCGTCA | AATGTTAGAA | 8280 |
| 30 | GATACAAAGC | CAACAACATT | TTCTGAATTA | GTTCAAATCT | CAGGATTATC | TCATGGTAÇA | 8340 |
| | GATGTGTGGT | TAGGCAATGC | TCAAGAATTA | ATTAAAACCG- | GTATATGTGA | TTTATCAAGT | 8400 |
| | GTAATTGGTT | GTCGTGATGA | TATCATGGTT | TATTTAATGT | ATGCTGGTTT | AGAACCATCA | 8460 |
| 35 | ATGGCTTTTA | AAATAATGGA | GTCAGTACGT | AAAGGTAAAG | GTTTAACTGA | AGAAATGATT | 8520 |
| | GAAAÉGATGA | aagaaaatga | AGTGCCAGAT | TGGTATTTAG | ATTCATGTCT | TAAAATTAAG | 8580 |
| 40 | TACATGTTCC | CTAAAGCCCA | TGCAGCAGCA | TACGTTTTAA | TGGCAGTACG | TATCGCATAT | 8640 |
| • | TTCAAAGTAC | ATCATCCACT | TTATTACTAT | GCATCTTACT | TTACAATTCG | TGCGTCAGAC | 8700 |
| | TTTGATTTAA | TCACGATGAT | TAAAGATAAA | ACAAGCATTC | GAAATACTGT | AAAAGACATG | 8760 |
| 45 | TATTCTCGCT | ATATGGATCT | AGGTAAAÀAA | GAAAAAGACG | TATTAACAGT | CTTGGAAATT | 8820 |
| | atgaatgaaa | TGGCGCATCG | AGGTTATCGA | ATGCAACCGA | TTAGTTTAGA | AAAGAGTCAG | 8880 |
| | GCGTTCGAAT | TTATCATTGA | AGGCGATACA | CTTATTCCGC | CGTTCATATC | AGTGCCTGGG | 8940 |
| 50 | CTTGGCGAAA | ACGTTGCGAA | ACGAATTGTT | GAAGCTCGTG | ACGATGGCCC | ATTTTTATCA | 9000 |
| | AAAGAAGATT | TAAACAAAAA | AGCTGGATTA | TCTCAGAAAA | TTATTGAGTA | TTTAGATGAG | 9060 |

| | GAAATAATCA | AGGTATTTAT | TTAATGCGTA | TGGCGTAGTC | AAAGAAATAC | AAAATTGTTG | 9180 |
|----|-------------|------------|------------|-------------|------------|------------|-------|
| | CTGGACACAA | AATTATGCCC | GTATTTCTTT | TCAATGTCTT | ACGAGTCTAT | TCAAATGTAA | 9240 |
| 5 | TGGTGAAATA | AAGGAACAAA | CTTTTACAAG | AATCTCTGAT | TAATAGTGAA | GTCATTTGTT | 9300 |
| | TCAAGCATAA | ACTTATGCTA | TAATTAAGTT | GCTTAAAAAT | TAGTGAACTC | AGGCAGAAGA | 9360 |
| | GTGGGAGATT | CCCGCTCTTT | TCTATTTGCC | AAAAAGGGAG | GCCTGTATGA | GTAAAATTAC | 9420 |
| 10 | AGAACAAGTA | GAAGTGATTG | TTAAACCAAT | TATGGAAGAC | TTGAATTTTG | AACTTGTAGA | 9480 |
| | CGTTGAATAT | GTCAAAGAGG | GTAGAGATCA | TTTTCTTAGA | ATCTCTATTG | ATAAAGAAGG | 9540 |
| 15 | TGGCGTAGAT | TTAAATGATT | GTACGCTAGC | TTCTGAAAAA | ATAAGTGAAG | CTATGGATGC | 9600 |
| | AAATGATCCT | ATTCCTGAAA | TGTATTATTT | AGACGTAGCG | TCACCTGGTG | CAGAACGTCC | 9660 |
| | AATTAAAAA | GAACAAGATT | TCCAAAATGC | AATAACTAAA | CCTGTATTTG | TTTCTTTATA | 9720 |
| 20 | TGTACCAATT | GAAGGTGAAA | AGGAATGGTT | AGGCATTTTA | CAAGAAGTCA | ATAATGAAAC | 9780 |
| | AATTGTAGTA | CAAGTTAAAA | TCAAAGCAAG | AACGAAAGAT | ATAGAGATAC | CGAGAGACAA | 9840 |
| | AATAGCAAAA | GCACGTCACG | CAGTTATGAT | TTAACGTGAT | GAGGAGGAAA | AAACGTGTCA | 9900 |
| 25 | AGTAATGAAT | TATTATTAGC | TACTGAGTAT | TTAGAAAAAG | AAAAGAAGAT | TCCTAGAGCA | 9960 |
| | GTATTAATTG | ATGCTATTGA | AGCAGCTTTA | ATTACTGCAT | ACAAAAAGAA | TTATGATAGT | 10020 |
| | GCAAGAAATG | TCCGTGTGGA | ATTAAATATG | GATCAAGGTA | CTTTCAAAGT | TATCGCTCGT | 10080 |
| 30 | AAAGATGTTG | TTGAAGAAGT | ATTTGACGAC | AGAGATGAAG | TGGATTTAAG | TACAGCGCTT | 10140 |
| | GTTAAAAACC | CTGCATATGA | AATTGGTGAT | ATATACGAÁG | AAGATGTAAC | ACCTAAAGAT | 10200 |
| | TTTGGTCGTG | TAGGTGCTCA | AGCAGCGAAA | CAAGCAGTAA | TGCAACGTCT | TCGTGATGCT | 10260 |
| 35 | GAACGTGAAA | TTTTATTTGA | AGAATTTATA | GACAAAGAAG | AAGACATACT | TACTGGAATT | 10320 |
| | ATTGÂCCGTG | TTGACCATCG | TTATGTATAT | GTGAATTTAG | GTCGTATCGA | AGCTGTTTTA | 10380 |
| 40 | TCTGAAGCAG | AAAGAAGTCC | TAACGAAAAA | TATATTCCTA | ACGAACGTAT | CAAAGTATAT | 10440 |
| | GTTAACAAAG- | TGGAACAAAC | GACAAAAGGT | -CCTCAAATCT | ATGTTTCTCG | TAGCCATCCA | 10500 |
| | GGTTTATTAA | AACGTTTATT | TGAACAAGAA | GTTCCAGAAA | TTTACGATGG | TACTGTAATT | 10560 |
| 45 | GTTAAATCAG | TAGCACGTGA | AGCTGGCGAT | CGCTCTAAAA | TTAGTGTCTT | CTCTGAAAAC | 10620 |
| | AATGATATAG | ATGCTGTTGG | TGCATGTGTT | GGTGCTAAAG | GCGCACGTGT | TGAAGCTGTT | 10680 |
| | GTTGAAGAGC | TAGGTGGTGA | AAAAATCGAC | ATCGTTCAAT | GGAATGAAGA | TCCAAAAGTA | 10740 |
| 50 | TTTGTAAAAA | ATGCTTTAAG | CCCTTCTCAA | GTTTTAGAAG | TTATTGTTGA | TGAAACAAAT | 10800 |
| | CAATCTACAG | TAGTTGTTGT | TCCTGATTAT | CAATTGTCAT | TAGCGATTGG | TAAAAGAGGA | 10860 |

| | GATGCGCGTG | AAGCGGGTAT | CTATCCAGTA | GTTGAAGCTG | AAAAAGTAAC | TGAAGAAGAT | 10980 |
|----|------------|------------|--------------------|------------|------------|------------|-------|
| | GTTGCTTTAG | AAGATGCTGA | CACAACAGAA | TCAACCGAAG | AGGTAAATGA | TGTTTCAGTT | 11040 |
| 5 | GAAACAAATG | TAGAGAAAGA | ATCTGAATAA | TAGGTTGGAG | TGAAGTATCT | ATGAAAAGA | 11100 |
| | AAAAAATTCC | GATGCGAAAA | TGTATTCTTT | CAAATGAAAT | GCATCCCAAA | AAAGATATGA | 11160 |
| | TTCGTGTTGT | TGTTAATAAA | GAAGGCGAAA | TCTTTGCGGA | TGTTACTGGA | AAGAAACAAG | 11220 |
| 0 | GCCGTGGCGC | ATATGTTTCT | AAAGATGTTG | CTATGGTTGA | AAAAGCACAA | CAAAAAGAAA | 11280 |
| | TTTTAGAAAA | ATATTTTAAA | GCATCTAAAG. | AGCAATTGGA | TCCTGTTTAC | AAAGAAATTA | 11340 |
| 5 | TTAGATTAAT | TTATAGAGAA | GAGATCCCAA | AATGAGTATA | GATCAAATAT | TAAACTTTTT | 11400 |
| 3 | AGGATTAGCA | ATGAGAGCTG | GTAAAGTAAA | AACAGGTGAA | TCAGTCATTG | TTAATGAGAT | 11460 |
| | TAAAAAAGGA | AATTTGAAGC | TCGTTATTGT | TGCAAATGAT | GCGTCTGATA | ATACAGCTAA | 11520 |
| 20 | ATTAATTACA | GATAAATGTA | AGAGTTACAA | AGTTCCATTC | AGAAAGTTTG | GAAATCGAAA | 11580 |
| | TGAATTGGGA | ATAGCACTTG | GAAAAGGTGA | GCGTGTTAAT | GTAGGGATTA | CTGACCCAGG | 11640 |
| | CTTTGCTAAA | AAGTTGCTAT | CAATGATAGA | TGAATATCAT | AAGGAGTGAT | TATATGAGTA | 11700 |
| 25 | AACAAAGAAT | TTACGAATAT | GCGAAAGAAT | TAAATCTAAA | GAGTAAAGAG | ATTATAGATG | 11760 |
| | agttaaaaag | CATGAATATT | GAGGTTTCAA | ATCATATGCA | AGCTTTGGAA | GATGACCAAA | 11820 |
| | TTAAAGCATT | AGATAAAAAG | TTCAAAAAAG | AACAAAAGAA | CGACAATAAA | CAAAGCACTC | 11880 |
| 80 | AAAATAATCA | CCAAAAATCA | AACAATCAAA | ACCAAAATAA | AGGGCMACAA | AAAGATAACA | 11940 |
| | AAAAGAATCm | ACAACAAAAT | AATAAAGGCA | ACAAAGGCAA | TAAAAAGAAT | AATAGAAATa | 12000 |
| | ataagaaaaa | TAACAAGAAT | AATAAACCAC | AAAATCAACC | AGCTGCTCCA | AAAGAAATAC | 12060 |
| 35 | CATCAAAAGT | GACATATCAA | GAAGGTATTA | CAGTAGGCGA | ATTTGCGGAT | AAATTAAATG | 12120 |
| | TTGĄĄTCATC | AGAAATTATC | AAAAAATTAT | TCTTACTTGG | TATTGTTGCT | AATATCAATC | 12180 |
| | AATCATTAAA | TCAAGAAACA | ATCGAATTAA | TTGCCGATGA | TTATGGCGTT | GAGGTTGAAG | 12240 |
| 10 | AAGAAGTTGT | GATTAATGAA | GAAGACTTAT | CAATCTATTT | CGAAGACGAA | AAAGATGATC | 12300 |
| | CAGAGGCAAT | TGAGAGACCA | GCAGTTGTAA | CAATTATGGG | ACATGTTGAC | CATGGTAAAA | 12360 |
| 15 | CGACTTTATT | AGATTCAATT | CGTCATACAA | AAGTTACAGC | AGGTGAAGCA | GGCGGAATCA | 12420 |
| | CTCAACATAT | TGGTGCATAT | CAAATTGAAA | ACGATGGCAA | AAAAATCACT | TTCTTAGATA | 12480 |
| | CACCGGGACA | TGCTGCATTT | ACAACGATGC | GTGCGCGTGG | TGCaCAAGTA | ACAGATATTA | 12540 |
| 50 | CTATTTTAGT | AGTAGCAGCT | GACGATGGTG | TTATGCCACA | AACAATTGAA | GCAATTAACC | 12600 |
| | ATCCTABAGA | ACCACAACTA | CC & A TETTA TETTS | TTGCAGTAAA | тававттсат | BABCCBACTT | 12660 |

| | GCGGCGAAAC | AATTETCGTC | CACTITCIGO | ATTAAGTGGT | GATGGTATCG | ACGATTTATT | 1278 |
|----|-------------|------------|------------|--------------|------------|------------|-------|
| | AGAAATGATA | GGATTAGTTG | CAGAAGTTCA | AGAACTTAAA | GCAAATCCTA | AAAACCGTGC | 1284 |
| 5 | TGTTGGTACA | GTTATCGAAG | CTGAATTAGA | TAAATCACGT | GGTCCTTCTG | CATCATTATT | 1290 |
| | AGTACAAAAC | GGTACATTAA | ATGTTGGTGA | TGCGATTGTA | GTTGGTAATA | CTTACGGCCG | 12960 |
| 10 | TATCCGTGCA | ATGGTTAATG | ACTTAGGTCA | AAGAATCAAA | ACGGCTGGTC | CATCAACGCC | 13020 |
| ,, | TGTTGAAATT | ACAGGTATTA | ATGATGTGCC | ACAAGCTGGG | GATCGCTTTG | TTGTATTTAG | 13080 |
| | TGATGAAAAA | CAAGCTCGTC | GTATTGGTGA | ATCAAGACAC | GAAGCTAGCA | TTATACAACA | 13140 |
| 15 | ACGTCAAGAA | AGTAAAAATG | TTTCATTAGA | TAACCTGTTT | GAACAAATGA | AACAAGGTGA | 13200 |
| | AATGAAAGAT | TTAAACGTTA | TTATTAAAGG | TGATGTTCAA | GGTTCTGTTG | AAGCTTTAGC | 13260 |
| | TGCATCATTA | ATGAAAATTG | ATGTTGAAGG | CGTAAATGTT | CGTATCATTC | ATACAGCGGT | 13320 |
| 20 | TGGTGCAATT | AATGAGTCAG | ACGTGACACT | TGCTAATGCC | TCAAATGGTA | TTATCATTGG | 13380 |
| | TTTCAATGTT | CGTCCAGACA | GTGGTGCAAA | ACGTGCTGCA | GAAGCTGAAA | ATGTTGATAT | 13440 |
| | GCGTTTACAC | AGAGTTATTT | ATAATGTTAT | CGAAGAAATT | GAATCAGCGA | TGAAAGGTTT | 13500 |
| 25 | ACTTGATCCA | GAATTTGAAG | AACAAGTTAT | CGGACAAGCT | GAAGTTCGTC | AAACATTCAA | 13560 |
| | AGTTTCTAAA | GTTGGTACTA | TTGCTGGATG | TTATGTTACT | GAAGGTAAAA | TTACGCGAAA | 13620 |
| | TGCTGGTGTA | CGTATTATTC | GTGATGGTAT | TGTTCAATAT | GAAGGCGAAT | TAGATACACT | 13680 |
| 30 | TAAACGTTTC | AAAGATGATG | CTAAGGAAGT | TGCAAAAGGT | TATGAATGTG | GTATTACAAy | 13740 |
| | TGAAAACTAC | AATGACCTTA | AAGAAGGCGA | TGTTATCGAA | GCATTTGAAA | TGGTTGAAAT | 13800 |
| • | TAAGCGTTAA | TTAAATAAAT | TACAAGCTAA | AAGTATAGTT | AAGATTGATA | TGCTCCCTAT | 13860 |
| 35 | AAATATTGCA | CTTTTTAAGT | GTCTACTTTA | TAGGGAGCAT | ATTTGATACT | AGCTTTTGGT | 13920 |
| | TITTIATTAG | AATAGATTAC | CTATTAAAAG | TTACGTTATA | TGGACATGAT | TTTGTATAAA | 13980 |
| 10 | ATTTTGTGGT | GGCCTAGAAT | GATTTTTAAT | GACAAAATAT | AATGTCGACT | ATTATTGGAA | 14040 |
| | AATTTTCTGT- | TGaAATGCCT | ATCTTACGGC | -AAACTTTATT- | TGATTTTATA | GGCTTAATTT | 14100 |
| | AATAAAATTA | CGTGTGAGCT | AAAATAATTG | TTTAAGCATT | GTTACACTAA | AAAATGCAAA | 14160 |
| 15 | TAACAATTGA | ACTTAAAGAT | AAAGAGGTGA | CAAGAATGAG | CAGTATGAGA | GCAGAGCGTG | 14220 |
| | TTGGTGAACA | AATGAAGAAG | GAATTAATGG | ATATCATCAA | CAATAAAGTC | AAAGATCCTC | 14280 |
| | GAGTTGGTTT | TATTACAATT | ACAGATGTTG | TTTTAACAAA | TGATTTATCG | CAGGCTAAAG | 14340 |
| 50 | TATTTTTAAC | TGTATTAGGT | AACGATAAAG | AAGTAGAAAA | TACATTTAAA | GCACTTGATA | 14400 |
| | AAGCAAAAGG | CTTCATTAAG | TCTGAATTAG | GTTCTAGAAT | GCGATTACCT | ATTATICCCC | 14460 |

| | AAGATTTACA | CAAACAAGAT | AGATAATTTA | GTGTTAGGTA | TCTGGAAAAT | GTTTGATAAT | 14580 |
|----|------------|------------|----------------|------------|------------|------------------|-------|
| | TTCTTAATAT | CGGTATATTA | ACATTAAACA | GTTAATACAT | AGATGTGTAG | AAATAGTTAA | 14640 |
| 5 | CATTTTCCAG | TTTTTTTATG | ATTAAATTA | GTTGATACGC | TATTAAAATA | TATTTTAAAA | 14700 |
| | AAGAAGGTGA | CTATATGTAT | AATGGGATAT | TACCAGTATA | TAAAGAGCGC | GGTTTAACAA | 14760 |
| | GTCATGACGT | TGTATTCAAA | TTGCGTAAAA | TATTAAAAAC | тааааааата | GGTCACACGG | 14820 |
| 0 | GTACGCTTGA | TCCCGAAGTT | GCAGGCGTGT | TACCGGTATG | TATAGGTAAT | GCAACGAGAG | 14880 |
| | TTAGTGATTA | TGTTATGGAT | ATGGGCAAAG | CTTATGAAGC | AACTGTATCG | ATAGGAAGAA | 14940 |
| 5 | GTACAACGAC | TGAAGATCAA | ACGGGTGATA | CATTGGAAAC | AAAAGGTGTA | CACTCAGCAG | 15000 |
| | ATTTTAATAA | GGACGATATT | GACCGATTGT | TAGAAAGTTT | TAAAGGTATC | ATTGAACAAA | 15060 |
| | TTCCGCCGAT | GTACTCATCC | GTCAAAGTAA | ATGGTAAAAA | attatatgaa | TATGCGCGTA | 15120 |
| 0 | ATAATGAAAC | agttgaaaga | CCAAAGCGTA | AAGTLAATAT | TAAAGACATT | GGGCGTATAT | 15180 |
| | CTGAATTAGA | TTTTAAAGAA | AATGAGTGTC | ATTTTAAAAT | ACGCGTCATC | TGTGGTAAAG | 15240 |
| | GTACATATAT | TAGAACGCTA | GCAACTGATA | TTGGTGTGAA | ATTAGGCTTT | CCGGCACATA | 15300 |
| 5 | TGTCGAAATT | AACACGAATC | GAGTCTGGTG | GATTTGTGTT | GAAAGATAGC | CTTACATTAG | 15360 |
| | AACAAATAAA | AGAACTTCAT | GAGCAGGATT | CATTGCAAAA | TAAATTGTTT | CCTTTAGAAT | 15420 |
| | ATGGATTAAA | GGGTTTGCCA | AGCATTAAAA | TTAAAGATTC | GCACATAAAA | AAACGTATTT | 15480 |
| 10 | TAAATGGGCA | GAAATTTAAT | AAAAATGAAT | TTGATAACAA | AATTAAAGAC | CAAATTGTAT | 15540 |
| | TTATTGATGA | TGATTCAGAA | AAAGTATTAG | CAATTTATAT | GGTACACCCT | ACAAAAGAAT | 15600 |
| | CAGAAATTAA | ACCTAAAAAA | GTCTTTAATT | AAAGGAGATA | GAATTTATGA | AAGTCATAGA | 15660 |
| 15 | AGTGACACAT | CCTATACAAT | CTAAACAGTA | TATTACAGAG | GATGTTGCAA | TGGCATTCGG | 15720 |
| | ATTITICGAT | GGCATGCATA | AAGGTCATGA | CAAAGTCTTT | GATATATTAA | ACGAAATAGC | 15780 |
| 10 | TGAGGCACGC | AGTTTAAAAA | AAGCGGTGAT | GACATTTGAT | CCGCATCCGT | CTGTCGTGTT | 15840 |
| 10 | GAATCCTAAA | AGAAAACGAA | CAACGTATTT | AACGCCACTT | TCAGATAAAA | TCGAAAAAAT | 15900 |
| | TAGCCAACAT | GATATTGATT | ATTGTATAGT | GGTTAATTTT | TCATCTAGGT | TTGCTAATGT | 15960 |
| 15 | GAGCGTAGAA | GATTTTGTTG | AAAATTATAT | AATTAAAAAT | AATGTAAAAG | AAGTCATTGC | 16020 |
| | TGGTTTTGAT | TTTACTTTTG | GTAAATTTGG | AAAAGGTAAT | ATGACTGTAC | TTCAAGAATA | 16080 |
| | TGATGCGTTT | AATACGACAA | TTGTGAGTAA | ACAAGAAATT | GAAAATGAAA | AAATTTCTAC | 16140 |
| 50 | AACTTCTATT | CGTCAAGATT | TAATCAATGG | TGAGTTGCAA | AAAGCGAATG | ATGCTTTAGG | 16200 |
| | ~~~~~~ | | 1-01-070071-07 | 0011000011 | ******** | G N N GWYN WYGGG | 16266 |

| | TGCTGTTAGT | ATTGAAATCG | GCACTGAAAA | TAAATTATAT | CGAGGGGTAG | CTAACATAGG | 16380 |
|----|------------|------------|------------|------------|------------|------------|-------|
| _ | TGTAAAGCCA | ACATTTCATG | ATCCTAACAA | AGCAGAAGTT | GTCATCGAAG | TGAATATCTT | 16440 |
| 5 | TGACTTTGAG | GATAATATTT | ATGGTGAACG | AGTGACCGTG | AATTGGCATC | ATTTCTTACG | 16500 |
| | TCCTGAGATT | AAATTTGATG | GTATCGACCC | ATTAGTTAAA | CAAATGAACG | ATGATAAATC | 16560 |
| 10 | GCGTGCTAAA | TATTTATTAG | CAGTTGATTT | TGGTGATGAA | GTAGCTTATA | ATATCTAGAG | 16620 |
| | TTGCGTATAG | tTATATAAAC | AATCTATACC | ACACCTTTTT | CTTAGTAGGT | CGAATCTCCA | 16680 |
| | ACGCCTAACT | CGGATTAAGG | AGTATTCAAA | CATTTTAAGG | AGGAAATTGA | TTATGGCAAT | 16740 |
| 15 | TTCACAAGAA | CGTAAAAACG | AAATCATTAA | AGAATACCGT | GTACACGAAA | CTGATACTGG | 16800 |
| | TTCACCAGAA | GTACAAATCG | CTGTACTTAC | TGCAGAAATC | AACGCAGTAA | ACGAACACTT | 16860 |
| | ACGTACACAC | AAAAAAGACC | ACCATTCACG | TCGTGGATTA | TTAAAAATGG | TAGGTCGTCG | 16920 |
| 20 | TAGCATTTAT | TAAACTACTT | ACGTAGTAAA | GATATTCAAC | GTTACCGTGA | ATTAATTAAA | 16980 |
| | TCACTTGGTA | TCCGTCGTTA | ATCTTAATAT | AACGTCTTTG | AGGTTGGGGC | ATATTTATGT | 17040 |
| | TCCAACCTTA | ATTTATATTA | AAAAAGCTTT | TTACAAATAT | TAACATTTAT | TATATGTTAA | 17100 |
| 25 | GCTAATATTG | AGTGAATAAT | AAGGTTACAA | TGAGATAAAG | ATGATATAAG | TACACCTAGA | 17160 |
| | GTAATAATCA | AGATATTAAA | AATAAAGTAT | GTTTTTTTAA | AAAATATAAC | TTATATTTAT | 17220 |
| | ACTGATAAGG | GTGGGACGAT | AAGTCTATTT | TGTAAATAAT | AGATGGATAT | CCCGCTCTCT | 17280 |
| 30 | TTTTTTCCAA | TTCAATATTT | TATAACTAAT | ATTAAAATAC | GATAATAAAT | GATATGATAT | 17340 |
| | AACTATTAGA | TTCAAGAGAG | GAGATTTATA | ATGTCTCAAG | AAAAGAAAGT | TTTTAAAACT | 17400 |
| 35 | GAATGGGCAG | GAAGATCTTT | AACGATTGAA | ACAGGGCAAT | TAGCTAAACA | AGCAAATGGC | 17460 |
| 55 | GCTGTATTGG | TTCGTTATGG | AGATACAGTC | GTGTTATCGA | CGGCAACTGC | ATCAAAAGAA | 17520 |
| | CCTCGTGATG | GAGATITCTT | CCCATTAACA | GTGAACTATG | AAGAAAAAAT | GTACGCTGCG | 17580 |
| 40 | GGTAAAATTC | CTGGTGGATT | TAAAAAGAGA | GAAGGACGTC | CTGGTGACGA | TGCAACATTA | 17640 |
| | ACTGCGCGAT | TAATTGATAG | ACCAATTAGA | CCTTTATTCC | CTAAAGGATA | TAAGCATGAT | 17700 |
| | GTTCAAATTA | TGAACATGGT | ATTAAGTGCA | GATCCTGATT | GTTCACCACA | AATGGCTGCA | 17760 |
| 45 | ATGATTGGTT | CATCTATGGC | GCTTAGTGTG | TCGGATATTC | CATTCCAAGG | GCCAATCGCC | 17820 |
| | GGTGTAAATG | TGGGTTATAT | TGACGGTAAA | TATATCATTA | ACCCAACAGT | AGAAGAAAAA | 17880 |
| | GAAGTTTCTC | GTTTAGACCT | TGAAGTAGCT | GGTCATAAAG | ATGCGGTAAA | CATGGTAGAG | 17940 |
| 50 | GCAGGCGCTA | GTGAGATTAC | TGAACAAGAA | ATGTTAGAGG | CGATTTTCTT | TGGTCATGAA | 18000 |
| | GAGATTCAAC | GTTTAGTTGA | TTTCCAACAA | CAAATCGTCG | ACCACATTCA | ACCIGITAAA | 18060 |

| | GAAGAAAAG | GACTTAAAGA | AACAGTTTTA | ACATTIGATA | AACAACAACG | AGATGAAAAT | 18180 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | CTTGATAACT | TAAAAGAAGA | AATCGTCAAT | GAATTTATCG | ATGAAGAAGA | TCCAGAGAAT | 18240 |
| 5 | GAATTACTTA | TTAAAGAAGT | TTATGCAATT | TTAAATGAAT | TAGTGAAAGA | AGAAGTTCGA | 18300 |
| | CGTTTAATTG | CAGATGAAAA | AATTAGACCA | GACGGCCGTA | AACCTGATGA | AATCCGTCCA | 18360 |
| 10 | TTAGATTCTG | AAGTTGGTAT | TTTACCTAGA | ACGCATGGTT | CAGGTCTATT | TACACGTGGT | 18420 |
| | CAGACTCAAG | CACTTTCAGT | TTTAACATTA | GGTGCTTTAG | GCGATTATCA | ATTAATTGAT | 18480 |
| | GGTTTAGGAC | CTGAAGAAGA | AAAAAGATTC | ATGCATCATT | ACAACTTCCC | GAATTTTTCA | 18540 |
| 15 | GTAGGTGAAA | CTGGTCCAGT | ACGTGCGCCA | GGTCGTCGTG | AAATTGGACA | TGGTGCGTTA | 18600 |
| | GGTGAAAGAG | CATTAAAATA | TATTATTCCT | GATACTGCTG | ATTTCCCATA | TACAATTCGT | 18660 |
| | ATTGTAAGTG | AGGTACTTGA | ATCAAATGGT | TCATCATCTC | AAGCGTCAAT | TTGTGGATCA | 18720 |
| 20 | ACATTAGCAT | TAATGGATGC | GGGCGTACCG | ATTAAAGCAC | CAGTTGCTGG | TATTGCTATG | 18780 |
| | GGCCTTGTTA | CACGTGAAGA | TAGCTATACG | ATTTTAACTG | ATATCCAAGG | TATGGAAGAT | 18840 |
| 25 | GCATTAGGTG | ATATGGACTT | TAAAGTCGCT | GGTACTAAAG | AAGGTATTAC | AGCAATCCAA | 18900 |
| 25 | ATGGATATTA | AAATTGACGG | TTTAACGCGT | GAAATTATCG | AAGAGGCTCT | AGAACAAGCG | 18960 |
| | AGACGTGGTC | GTTTAGAAAT | AATGAATCAT | ATGTTACAAA | CAATTGATCA | ACCACGTACT | 19020 |
| 30 | GAATTAAGTG | cTTACGCGCC | AAAAGTTGTA | ACTATGACAA | TTAAACCAGA | TAAGATTAGA | 19080 |
| | GATGTTATCG | GACCTGGTGG | TAAAAAAATT | AACGAAATTA | TTGATGAAAC | AGGTGTTAAA | 19140 |
| | TTAGATATTG | AACAAGATGG | TACTATCTTT | ATTGGTGCTG | TTGATCAAGC | TATGATAAAT | 19200 |
| 35 | CGTGCTCGTG | AAATCATTGA | GGAAATTACA | CGTGAAGCGG | AAGTAGGTCA | AACTTATCAA | 19260 |
| | GCCACTGTTA | AACGTATTGA | AAAATACGGT | GCGTTTGTAG | GCCTATTCCC | AGGTAAAGAT | 19320 |
| | eccijeciic | ACATTTCACA | AATTTCAAAA | AATAGAATTG | AAAAAGTGGA | AGATGTATTA | 19380 |
| 40 | AAAATCGGTG | ACACAATTGA | AGTTAAGATT | ACTGAAATTG | ATAAACAAGG | TCGAGTAÀAT | 19440 |
| | GCTTCACATA | GAGCATTAGA | AGAATAATAT | TTAAAGTCAT | ATGACGACAA | TGTATCGTCA | 19500 |
| 45 | TGTGATTTTT | TTATGCCACT | TTTTACGAAG | TGACCCGTTT | TGAATTTGTT | GTATTGAACA | 19560 |
| | TTTTAAAACG | CTTTATTATT | TTGTGTGCAA | CTGTTAATTA | TCCTGTATGT | ATAGTGATTA | 19620 |
| | ATAGTGTACA | TCAAGTGTTT | TTTAACTTAT | AATGAATAGT | GAGTTTATAT | ATGGACGGGT | 1968 |
| 50 | AACAAATTTA | GGAGGTAAGA | TTTTGAGTTT | AATAAAGAAA | AAGAATAAAG | ATATTCGCAT | 1974 |
| | TATACCATTA | GGCGGTGTTG | GCGAAATTGC | TAAAAATATG | TATATCGTTG | AAGTAGACGA | 1980 |
| | TGAAATGTTT | ATGTTAGATG | CTGGACTTAT | GTTTCCAGAA | GACGAAATGC | TAGGTATTGA | 1986 |

| | CCTTACACAC | GGACATGAGC | ACGCGATTGG | TGCAGTGAGT | TATGTTTTAG | AACAATTAGA | 19980 |
|----|--------------|-------------|------------|------------|------------|------------|-------|
| | TGCACCAGTA | TATGGATCTA | AATTGACAAT | AGCGTTAATI | AAAGAAAATA | TGAAAGCCCG | 20040 |
| 5 | TAATATTGAT | AAAAAAGTTC | GCTACTATAC | AGTTAATAAT | GATTCAATTA | TGAGATTCAA | 20100 |
| | AAACGTGAAT | ATTAGTTTCT | TTAATACGAC | ACACAGTATT | CCTGATAGTT | TAGGTGTTTG | 20160 |
| 10 | TATTCACACT | TCATATGGTG | CCATTGTGTA | TACAGGTGAA | TTTAAGTTTG | ACCAAAGTTT | 20220 |
| | ACATGGACAT | TATGCACCAG | ATATTAAACG | TATGGCAGAG | ATTGGTGAAG | AAGGCGTATT | 20280 |
| | TGTCTTAATC | AGTGATTCTA | CTGAGGCAGA | GAAACCTGGA | TATAATACTC | CGGAAAATGT | 20340 |
| 15 | GATTGAACAT | CATATGTATG | ATGCTTTTGC | AAAAGTGCGA | GGTCGCTTGA | TAGTTTCATG | 20400 |
| | TTATGCTTCG | AACTTTATAC | GTATTCAGCA | AGTTTTAAAT | ATTGCTAGCA | AGCTAAATCG | 20460 |
| | TAAAGTGTCA | TTTTTAGGAA | GATCACTTGA | AAGTTCATTT | AATATTGCTC | GTAAAATGGG | 20520 |
| 20 | GTATTTCGAC | ATTCCTAAAG | ATTTGCTAAT | TCCTATAACA | GAAGTTGATA | ATTATCCTAA | 20580 |
| | AAATGAAGTG | ATAATTATAG | CTACTGGTAT | GCAAGGAGAA | CCTGTAGAAG | CCTTAAGTCA | 20640 |
| | AATGGCGCAA | CATAAGCATA | AAATTATGAA | TATCGAAGAA | GGCGATTCTG | TATTTTTAGC | 20700 |
| ?5 | AATTACGGCT | TCTGCTAATA | TGGAAGTTAT | CATTGCGAAT | ACATTAAATG | AGCtTgTtAC | 20760 |
| | GnCTGGCGCA | CATATTATTC | CAAATAACAA | AAAGATTCAT | GCTTCAAGTC | ATGGTTGCAT | 20820 |
| 30 | GGAAGAATTA | aaaatgatga | TTAATATTAT | GAAACCTGAA | TACTTTATTC | CTGTACAAGG | 20880 |
| | TGAATTTAAA | ATGCAGATAG | CACATGCGAA | GCTAGCAGCT | GAAGCAGGTG | TTGCACCAGA | 20940 |
| | AAAGATTTTC | CTTGTGGAAA | AAGGAGATGT | CATTAATTAC | AACGGTAAAG | ATATGATATT | 21000 |
| 35 | AAATGAAAAG | GTAAATTCAG | GAAATATTTT | AATAGATGGC | ATTGGTATTG | GGGATGTAGG | 21060 |
| | AAATATCGTG | TTGAGAGACC | GTCATCTTTT | AGCAGAAGAT | GGTATCTTTA | TTGCTGTTGT | 21120 |
| | AACGTTAGAT | ССТАААААТА | GACGTATAGC | TGCGGGACCT | GAAATTCAAT | CTCGTGGGTT | 21180 |
| | TGTATATGTA | egtgaaagtg- | AAGACTTATT | ACGTGAAGCA | GAAGAGAAAG | TACGTGAAAT | 21240 |
| | AGTAGAGGCT | GGTTTACAAG | AAAAACGCAT | AGAATGGTCT | GAAATTAAAC | AAAATATGCG | 21300 |
| 15 | TGATCAAATT . | AGTAAACTAT | TATTCGAAAG | TACAAAACGT | CGTCCTATGA | TTATTCCAGT | 21360 |
| | AATTTCTGAA . | ATTTAATCAA | AAAGTCATTA | ACATAAAAGA | GGTCAGAACA | AGTCACTGAA | 21420 |
| | ATATAATGGT ' | TGTCATGGAC | AATTTACTTA | TATTTTATGA | TAGTCAATTG | AAGGGGTAAC | 21480 |
| 0 | GATTAATCTG | TTATCTTAAG | TAAATTGATA | CATAGATGAT | ATTGTTCTAA | CCTCTTTCAT | 21540 |
| | CGTCTGTTTG | GACTACATAT | TCTAAACATC | AAATAGGAAA | TTATATATAA | TAACGTCGTT | 21600 |
| | TTAACTAAGG | CAACATAAGG | AGGTGCGTCA | ATTGGCACAA | GCAAAAAAGA | AATCGACAGC | 21660 |

GATACGTTAT GTCATAGCTA TTTTAGTAGT TGTATTAATG GTGTTGGGTG TTTTCCAATT 21780 AGGAATAATA GGTCGTCTAA TTGACAGCTT CTTTAATTAT TTATTTGGGT ACAGTAGATA 21840 TTTAACATAT ATTTTAGTAC TCTTAGCAAC TGGTTTTATT ACATACTCTA AACGTATTCC 21900 TAMAACTAGA CGAACGGCTG GTTCGATTGT ATTGCAAATT GCATTGCTAT TTGTATCACA 21960 GTTAGTTTT CATTITAATA GTGGTATCAA AGCTGAAAGA GAACCTGTAC TTTCTTATGT 22020 GTATCAGTCA TACCAACACA GTCATTTCCC AAATTTTGGT GGCGGTGTAT TAGGCTTTTA 22080 TTTATTAGAG TTAAGCGTAC CTTTAATTTC ATTATTTGGT GTATGTATTA TTACTATTTT 22140 ATTATTATGC TCAAGTGTTA TTTTATTAAC AAACCATCAA CATCGTGAAG TTGCAAAAGT 22200 TGCACTGGAA AATATAAAAG CTTGGTTTGG TTCATTTAAT GAA 22243

(2) INFORMATION FOR SEQ ID NO: 165:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5510 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

25

30

35

40

45

50

10

15

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 165:

TTATTAAThA TTAATATTT TATTTTTAAA AATAAAGCGA GGAGCTATCA ATGGAACAAA 60 TTACTTCTGC ACAAAATAAT AGAATTAAAC AAGCGAACAA GCTAAAAMAG AAACGTGAGA 120 GGGATAAAAC TGGATTAGCT TTAATTGAAG GTGTGCATTT AATTGAAGAA GCTTATCAAA 180 GTGGAATTGT AATTACACAA TTATTTGCAA TTGAACCGGC AAGATTAGAT CAGCAAATTA 240 WCGCATACGC GCAAGAAGTT TTTGAAATAA ACATGAAAGT TGCTGAATCT TTATCAGGTA 300 CAGTGACACC ACAAGGGTTT TTCGCAATCA TTGAGAAGCC GCATTATGAT ATTTCTAAAG 360 CACAACAAGT ATTGCTCATC GATCGTGTTC AAGATCCTGG AAATTTAGGC ACATTAATTA 420 GAACTGCGGA TGCTGCTGGA ATGGATGCTG TAATAATGGA GAAGGGTACG ACAGATCCTT 480 ATCAAGATAA AGTGTTGCGA GCGAGTCAAG GTAGTGTTTT CCATTTGCCA GTTATGACAC 540 AAGATCTCGA TACGTTTATT ACTCAATTTA ATGGTCCTGT TTATGGTACA GCACTTGAAA 600 ACGCAGTGGC ATACAAAGAA GTTACTTCAA GTGATTCTTT TGCATTACTA TTAGGTAATG 660 720 AGGGAGAGG TGTTAATCCT GAATTATTAG CACATACTAC ACAAAATTTA ATCATACCTA TTTATGGTAA AGCTGAAAGT TTAAATGTAG CGATTGCAGG TAGTATTTTA CTTTATCATT 780 TGAAAGGTTG ACCGTGTTGA AAGTTTTCCG ATATAATTAT AATTAATTGT TTAACAGAAC 840

| | ATAAATAATT GTTTTAGGGA GAATAATCGT GACTGCAAGT TATTCCAATT ATTTAAAGTC | 960 |
|-----|---|-------|
| | TTTTCACCTT TTTGGTTACT TAAAGAGATT TAAGTCGGAA AGACAATCCG TTATCAATAT | 1020 |
| 5 | TAAACAAGTG TATGCTTAGG CATAAATTTG GGTGGTACCA CGGAAATGAC TTTCGTCCCT | 1080 |
| | TATTTTTAA GAGGATGAAA GTCTTTTTT AGTTAAACAA CAAATATGAT AAATAGAAAA | 1140 |
| 10 | TGAATAGTTC GAATAGGGAG GTCAGTGACA TATGTCTGAA CAACAAACAA TGTCAGAGTT | 1200 |
| 10 | AAAACAACAA GCGCTTGTAG ATATTAATGA AGCAAATGAT GAACGTGCAC TGCAAGAAGT | 1260 |
| | TAAAGTGAAA TACTTAGGTA AAAAAGGGTC AGTTAGCGGA CTAATGAAAT TGATGAAGGA | 1320 |
| 15 | TTTGCCGAAT GAAGATAAAC CTGCGTTTGG TCAAAAAGTG AATGAATTGC GTCAAACAAT | 1380 |
| | TCAAAATGAA TTAGATGAAA GACAACAGAT GTTAGTTAAA GAAAAATTAA ATAAGCCAAT | 1440 |
| | TGGCTGAAGA AACAATTGAT GTATCATTAC CAGGTCGTCA TATTGAAATC GGTTCAAAGC | 1500 |
| 20 | ATCCATTAAC ACGTACAATA GAAGAAATTG AAGACTTATT CTTAGGTTTA GGTTATGAAA | 1560 |
| | TTGTGAATGG ATATGAAGTT GAACAAGATC ATTATAACTT CGAAATGCTG AATTTACCTA | 1620 |
| | AATCACACCC TGCACGTGAT ATGCAAGATA GTTTCTATAT TACGGATGAA ATTTTATTAC | 1680 |
| ?5 | GTACGCATAC ATCACCAGTG CAGGCACGTA CGATGGAATC ACGTCATGGT CAAGGTCCAG | 1740 |
| | TTAAAATTAT TTGCCCTGGT AAAGTGTATC GTCGTGACTC TGATGATGCG ACACATAGTC | 1800 |
| 30 | ATCAATTTAC ACAAATCGAA GGATTAGTTG TTGATAAAAA CGTTAAAATG AGTGATTTGA | 1860 |
| | AAGGTACTTT AGAATTGTTA GCTAAGAAAT TATTTGGTGC TGATCGTGAA ATTCGTTTAC | 1920 |
| | GTCCAAGTTA CTTCCCATTC ACTGAACCTT CTGTAGAAGT TGATGTGTCA TGTTTTAAAT | 1980 |
| 15 | GTAAAGGAAA AGGTTGTAAT GTGTGTAAAC ACACAGGATG GATTGAAATT TTAGGTGCTG | 2040 |
| | GAATGGTACA TCCTAATGTA TTAGAAATGG CTGGTTTTGA TTCTTCAGAG TACTCTGGAT | 2100 |
| | TTGCATTTGG TATGGGACCA GACCGTATTG CAATGTTGAA ATATGGTATA GAAGATATTC | 2160 |
| 0 | GTCATTTCTA TACTAATGAT GTGAGATTTT TAGATCAATT TAAAGCGGTA GAAGATAGAG | _2220 |
| | GTGACATGTA ATGTTGATAT CAAATGAATG GTTGAAAGAA TATGTAACAA TCGATGATTC | 2280 |
| | TGTAAGTAAT TTGGCAGAAC GTATTACGCG CACAGGTATT GAAGTGGATG ATTTAATTGA | 2340 |
| 5 | CTACACAAAA GATATCAAAA ATTTAGTTGT CGGCTTCGTT AAGTCAAAAG AGAAACATCC | 2400 |
| | TGATGCTGAT AAATTAAATG TTTGCCAAGT TGATATCGGA GAAGACGAAC CTGTACAAAT | 2460 |
| o · | CGTTTGTGGT GCACCGAACG TTGATGCAGG ACAATATGTC ATTGTTGCTA AAGTAGGTGG | 2520 |
| - | CAGATTGCCT GGTGGTATTA AAATTAAGCG TGCCAAATTA CGCGGTGAAC GTTCAGAAGG | 2580 |
| | TATGATTTGT TCGTTACAAG AAATTGGTAT TTCAAGTAAC TATATACCGA AAAGTTTTGA | 2640 |

| | ATATTTAGAT | GATCAAGTAA | TGGAATTTGA | TTTAACGCCG | AATCGTGCAG | ATGCTTTAAG | 2760 |
|----|------------|------------|------------|------------|------------|------------|------|
| | TATGATAGGT | ACTGCTTATG | AAGTTGCAGC | ATTATATAAT | ACAAAAATGA | CTAAGCCAGA | 2820 |
| 5 | GACAACATCA | AATGAGCTTG | ATTTATCTGC | AAATGATGAA | CTGACTGTGA | CAATTGAAAA | 2880 |
| | TGAAGATAAA | GTACCATATT | ATAGTGCACG | TGTTGTTCAC | GACGTGACAA | TTGAACCCTC | 2940 |
| 10 | GCCAATTTGG | ATGCAAGCAC | GCTTAATAAA | AGCGGGTATA | CGTCCTATTA | ATAATGTTGT | 3000 |
| | TGACATTTCA | AATTATGTGT | TATTAGAATA | CGGTCAACCA | TTGCACATGT | TTGATCAAGA | 3060 |
| | TGCGATTGGT | TCACAACAAA | TTGTTGTTCG | TCAAGCTAAT | GAAGGCGAAA | AAATGACAAC | 3120 |
| 15 | ATTAGATGAT | ACAGAACGTG | AATTATTAAC | GAGCGATATT | GTCATTACTA | ATGGACAAAC | 3180 |
| | TCCAATTGCA | TTAGCTGGTG | TTATGGGTGG | CGATTTTTCA | GAAGTTAAAG | AACAAACATC | 3240 |
| | AAATATAGTG | ATTGAAGGTG | CTATTTTGA | TCCAGTTTCA | ATTCGTCATA | CATCAAGACG | 3300 |
| 20 | TTTAAATTTA | CGCAGTGAAT | CATCTAGTCG | TTTTGAAAAA | GGAATAGCTA | CTGAATTTGT | 3360 |
| | AGATGAAGCA | GTCGACCGTG | CATGTTATTT | ATTACAÂACT | TATGCAAACG | GAAAAGTGCT | 3420 |
| | AAAAGATAGA | GTGTCTTCAG | GAGAACTTGG | TGCATTTATT | ACACCAATCG | ACATCACTGC | 3480 |
| 25 | TGATAAAATT | AATCGCACTA | TTGGATTTGA | TTTGTCACAA | AATGATATTG | TTACTATTTT | 3540 |
| | TAATCAACTA | GGGTTTGATA | CAGAAATAAA | TGATGATGTT | ATTACAGTGC | TAGTACCATC | 3600 |
| 30 | ACGTCGTAAA | GATATTACAA | TTAAAGAAGA | TTTAATTGAA | GAAGTTGCAC | GTATATATGG | 3660 |
| | ATACGACGAT | ATTCCATCAA | CGTTACCTGT | CTTCGATAAA | GTTACTAGTG | GTCAGCTAAC | 3720 |
| | TGATCGCCAA | TATAAAACTA | GAATGGTTAA | AGAAGTGTTA | GAAGGTGCTG | GATTAGACCa | 3780 |
| 35 | AGCTATTACG | TATTCGTTAG | TTTCTAAAGA | AGATGCTACT | GCaTTTTCGA | TGCAACAGCG | 3840 |
| | TCAAACAATT | GATTTATTGA | TGCCAATGAG | TGAAGCGCAT | GCGTCATTAC | GTCAAAGTTT | 3900 |
| | ATTÁCCACAT | TTAATCGAAG | CGGCATCATA | TAATGTGGCA | CGCAAAAATA | AAGATGTAAA | 3960 |
| 40 | ATTÄTTTGAA | ATCGGCAATG | TCTTCTTTGC | TAATGGAGAA | GGTGAACTAC | CAGATCAAGT | 4020 |
| | TGAATATTTA | AGTGGTATTT | TAACTGGAGA | TTATGTAGTC | AATCAATGGC | AAGGTAAGAA | 4080 |
| 46 | AGAAACGGTT | GATTTCTATT | TAGCAAAAGG | TGTCGTGGAT | CGAGTATCTG | AAAAGTTAAA | 4140 |
| 45 | TCTTGAATTT | AGTTATCGCC | GTGCTGATAT | TGaTGGATTA | CATCCAGGTC | GTACTGCTGA | 4200 |
| | AATCTTATTA | GAGAATAAAG | TTGTTGGTTT | TATTGGTGAA | TTACATCCAA | TATTAGCAGC | 4260 |
| 50 | TGATAATGAT | TTAAAACGTA | CGTATGTTTT | TGAGTTGAAT | TTTGATGCAT | TAATGGCTGT | 4320 |
| | GTCGGTAGGT | TACATTAATT | ACCAGCCAAT | TCCGAGATTC | CCAGGCATGT | CTCGTGACAT | 4380 |
| | TGCATTAGAA | GTAGATCAAA | ATATTCCAGC | AGCTGATTTA | TTATCAACGA | TTCATGCACA | 4440 |

| | AAAAGGTAAA AAATCAATTG CAATACGTTT AAATTATTTA GACACAGAAG AAACATTGAC | 4560 |
|----|---|--|
| | AGATGAGCGC GTTTCAAAAG TACAAGCGGA AATTGAAGCA GCATTAATTG AACAAGGTGC | 4620 |
| 5 | TGTTATTAGA TAATGATTTA AACCCCATGT ATAAGGATAT CTGAAGTAGA TTGATATCCC | 4680 |
| | TAACATGGGG TTTTATTTTT GGGTTCACCA ATTTGGTTCC AATGCATTTA AAAAGTCAAA | 4740 |
| 10 | GAGGAACAGC GGAATACAGA TGATGCTTCG CACAACTGCA TAAAAGCCTC TAATGATTAA | 4800 |
| 10 | AAATCAAAGA GGCTTTAAAA TTTTTTGGGC TTTTTCACGA TTTTTAAAAT GCTTTTTTGA | 4860 |
| | AATGGTATCT AAACGTGAAA GACCGTATTT TTTTATAATT TTGGCGGCGA TTACATCGAC | 4920 |
| 15 | TTTAGCACCG GCACCTTTAG GAATCGTCAT ATTAATATTT TTTGATATTT GATCCATATA | 4980 |
| | TGTAACAAAT GCGTATCGAG AAATTATGCT TGCCACTGCA ATGGCTAATG ACTTCGATTC | 5040 |
| | TCCTTTTGTT TCAAATTTTG TTTTCTTTGG AAGTGGTATA TCTGATAATG CGTAATGGCT | 5100 |
| 20 | ATACACTTCG CGTTTTGCGA ACTGATCAAT GACGATATAG TCTAATTGAG ACGAATCAAT | 5160 |
| | TTTTTCAAGT ACATTTTTGA TGGCTTCATT ATGAAGGGCA GCTTTCATTT TTACTTGAGT | 5220 |
| | CCAGCCTTTT GCTTGCTGAA TATTATATTT TTCATTGTGT AGTGTTAATA ATGAATGTGG | 5280 |
| 25 | TATGAAAGTA ACCAATTGCT CAGCAAGTTC TACAATTTTG GTATCGGTTA ATTTTTTTGA | 5340 |
| | ATCATCTACA CCCAAAGTTT TTAAAATAGG GACATGCTCT TTGGTAACGA AAGCAGCACA | 5400 |
| 30 | CACAGTCAAC GGACCAAAGT AATCGCCACT TCCAGCCTCA TCACTACCAA TACAGTTAAA | 5460 |
| 30 | TTGTTCATAC ATTABAGTTG TCCAGAAAAG AATTAGCCAT ATTTTCCTTT | 5510 |
| - | (2) INFORMATION FOR SEQ ID NO: 166: | |
| 35 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 9623 base pairs | |
| | (B) TYPE: nucleic acid | |
| | (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 40 | | |
| | (x1) SEQUENCE DESCRIPTION: SEQ ID NO: 166: | · ———————————————————————————————————— |
| | GNTTATACTT ATAAATTTTA CGGGGGTAAT ATAATACTLA TTTACCTGTA ATATATGATA | 60 |
| 45 | ATTCTTCAGC GGCAGCTGCG TTGATAGTTC TATGAGAAAT GATACCTAAT CCTTTAACAT | 120 |
| | TGGATTCTGA AATAACGATA GAACCATCAC TGTTAACTTT TTCAACAAAT GCTACATGAC | 180 |
| | CGTAATGTTG ATCTGCACCA AATTGTCCAG CCTCAAATAC AACAGCAGCA TGACGTTTTG | 240 |
| 50 | GTGTATGACT TACTTGATAA TCACGGTATT GAGCTCGATT ATTCCAATTA TGTGCATCAC | 300 |

55

360

CTAAATCACC TGAGATAGAT GTACCAAATT GTTTCATACG GTTATATACG TACCAAGTAC

| | ATGAATCATC | ATAATCCTTG | ATAGAACGTT | CATATTTATC | TAAATCTGGC | ATGCGTTCAT | 480 |
|----|-----------------|------------|------------|------------------|------------|------------|------|
| | CGTCAAACTG | AGTTAATTGA | TAGTGTTTAA | TAATACTGTT | TAATTTCTTA | GCATAGTTTG | 540 |
| 5 | GATCTGTAGC | ATATGTTTTA | GATAAGTGTG | ATGTTGCATC | TTTATAAGAA | TCGGCTTCCG | 600 |
| | ATTTCCATGT | TGGTTTATAA | ATTGTTCGAT | TGCCATCAAT | ACCATTTTTA | ATAAGGTCAG | 660 |
| •• | AGTAATCTTT | TAGTGATTCT | TTCGTGCTTG | GATATTTTCG | GAATCCAGCA | TTAATACTAT | 720 |
| 10 | ACAATTGATT | ACCATCAGCT | TCTAATGTGT | TAAAAGGAAC | AGAATTCCCT | TCaAAAGCAC | 780 |
| | CTTTGATACC | GAATAAATTA | TGGTTTGGTG | ACWTAGCTAA | AGCACTACGA | CCTGAGTCAG | 840 |
| 15 | ATTCTAAGAT | TGCTTGGGCA | ATCATGACAG | ACGCATAAAT | ATCGTTATCT | TGACCAATGC | 900 |
| | GATGTGCATC | TTTAGCAATT | GATTTGACAA | ATTGACGTGT | ATCTTTTGAG | TCAACAACGT | 960 |
| | TAAATTGTCC | GCTATCATCA | TTGTTAGATA | TACTAGGATC | TGTTTCGAAT | AATGATGTTG | 1020 |
| 20 | CACGTGTATC | CTTTTGATTA | ACATCGTTAT | TGAATGATTG | AGCAGGTTTA | GATTTATGTT | 1080 |
| | TCAATTCATC | TTGTGTTGGT | AACTGTGGAT | TCTTTGTATT | AGATTTTTCA | TTTTTGTCTT | 1140 |
| | TTTTAGATTG | AGATGCATAA | TCTTTTTGTG | TTTTCTTTGC | ATCTTCACTG | TATTGATCCA | 1200 |
| 25 | AAATAGAGTC | TAAAGCCGAA | TCTGACATTG | ATTGATTATC | TTTCGATGAA | GATTTTTGAT | 1260 |
| | TTGCTTTATC | GTCACTTGCT | GGTTGACTAT | TTGATTGATT | AGGTTGTGTT | GGCTTTGGCG | 1320 |
| 30 | AATTTGGTTG | CTTATTAGAT | GTACTTGGTT | TTGTATTGTT | TGATTTAGGT | GCTTTTTGAT | 1380 |
| | TGTCTGCTTT | ATCTTGTTTA | GATGATTGCG | TATCAGTGTC | ATTTTTGATG | CTATTGTCAC | 1440 |
| | TGTTTTTATT | CGAATCATTT | GTTGACTTTT | CGCCATTACG | AGGTTGTTCG | TAATCAGAAA | 1500 |
| 35 | TATCCGAATT | TAAATTGAAT | AAGTTTTGGA | TTAAAGTTGT | TAATGAGTAA | TTATCATCGT | 1560 |
| | ATTTATTTTT | GGTTAGCAAT | TGGTTTATAT | TGGTTTGTGG | TAAATTCTTA | TAAATAAAT | 1620 |
| | CAATGATATT | GTTAGAGTCT | GAAGTGCTGT | CGTCTATAGT | TTTAAATTTT | TTGTCGTTAT | 1680 |
| 40 | TGTĊTTGGTT | ACTTGTATTA | TTTTTGTCTG | CTTTATCAAT | ATCTTTACTT | GTAGTATCCT | 1740 |
| | TAGAAGTTTC | ATCGTCATTA | GATTTTTTTG | AATCATGAGA | TGTTGTCTTA | GCTGTAGTAT | 180 |
| | CTTTTTGAGG | TGTATCAGCA | TAAGCGGTAG | GTGAAaCTAA | AGTAGGTAAT | ACGAGCGTAG | 186 |
| 45 | TTGATAGCAA | ATAAATTAAA | ATTTTATTTT | TAGGCATATT | TCGTATTCTC | CCTTGAAAAA | 192 |
| | TATAATAATT | AAGTGTGATA | ATAAACTATG | ATTTGTTATA | ATTTATCGTA | TGCTGAAAAT | 198 |
| 50 | AGTTGATAGG | TATCAATCGA | CTAAATATCT | TCCAGTAAAT | TGATTATACT | AATTCACAAC | 204 |
| | GCAAAAATAA | ATTAATTTAC | ATATAAAAA | TAAAAAATAT | GAATAATTCC | TACATAGGAG | 210 |
| | TOTO 3 C 3 3 TO | **** | **** | TA A B A WOO A T | CTCNACNAC | TACCTAACAC | 216 |

| | TAACTTATGG | GCAATGTGGG | ATCCATATGG | CAACACGGGA | CACATCAAGG | TCGCAGTCGT | 2280 |
|----|--------------|------------|------------|------------|------------|------------|------|
| | TAATGAAGAT | AAAGGCGACA | CAATCAGAGG | GAAAAAAGTT | AATGTCGGTA | ATACGATGGT | 2340 |
| 5 | TAATACACTC | AAGAAAAATA | AAAGTTTTGA | TTGGCAGTTI | GTAAGTAGAG | AGAAAGCTGA | 2400 |
| | TCATGAGATA | AAAATGGGTA | AATATTTTGC | AGGTATTTAC | ATCCCATCTA | AGTTTACACA | 2460 |
| 10 | TGAAATTACA | GGGACACTAC | GTAAGCAGCC | TCAAAAAGCA | GATGTAGAAT | TTAAGGTGAA | 2520 |
| | TCAGAAGATT | AACGCTGTTG | CGTCTAAGCT | AACAGATACT | GGTTCGTCAG | TTGTCGTTGA | 2580 |
| | AAAAGCGAAT | GAACAATTTA | ATAAAACAGT | AACTCGAGCA | TTATTAGAAG | AAGCTAACAA | 2640 |
| 15 | AGCAGGTTTA | ACTATTGAAG | AAAATGTGCC | GACAATTAAC | AAGATAAAA | ATGCGGTATA | 2700 |
| | TTCAGCAGAT | AAAGCTTTAC | CTAAGATTAA | TGACTTTGCG | AATAAAATTG | TATATTTGAA | 2760 |
| | TAACCACCAA | GCGGATTTAG | ATAAATATGC | CAATGATTTT | AGAAAACTAG | GAAATTATAA | 2820 |
| 20 | AGGTGATATT | TTAGATGCTC | AGAAAAAATT | AAACGAaGTC | AATGGTGCTA | TTCCGCAACT | 2880 |
| | TAATGAAAAG | GCTAAGTTGA | TATTAGCTTT | AAATAATTAT | ATGCCGAAAA | TTGAAAAAGC | 2940 |
| | GTTAAATTTT | GCAGCTGATG | ACGTGCCAGC | GCAGTTCCCT | AAAATTAATC | AAGGACTTAA | 3000 |
| 25 | CATTGCGAGT | CAAGGTATTG | ATCAAGCTAA | TGGACAGTTA | AATGATGCCA | AAGGCTTCGT | 3060 |
| | CACACAAGTT | AGAAGTAGAG | TCGGTGATTA | TCAAGATGCA | ATTCGACGCG | CGCAAGATTT | 3120 |
| 30 | AAATCGAAGA | AACCAGCAAC | AGATTCCTCA | AAATAGCGCG | GCGAACAACG | AAACATCAAA | 3180 |
| | TAGTGCACCT | GCAGCTGGTA | ATGGTGTAGC | ATCAACGCCA | CCAAGTGCAC | CAAGTGGCGA | 3240 |
| | TACTGCACCA | AATAATAATG | TTACGCAAAA | TACCGCACCA | AATAGTAATA | ATGCGCCTGT | 3300 |
| 35 | ATCGACTACA | CCACAAAGTA | CAAGCGGGAA | AAAAGATGGT | CAAAGTTTTG | TAGATATAAC | 3360 |
| | AACAACACAA | GTCAGCACAG | CTAACGAGAA | CACACAAAAC | ATTACAGATA | AAGATGTTAA | 3420 |
| | ATCAATGGAA | GCGGCATTAA | CGGGCTCTTT | ATTATCATTA | TCAAATAATT | TAGATACCCA | 3480 |
| 10 | AGCGÄAAGCC | GCACAAAAAG | ATAGTCAGGC | ATTACGTAAT | ATTTCGTATG | GGATTTTAGC | 3540 |
| | ATCGGACAAG | CCTTCTGATT | TTAGAGAGTC | TTTAGATAAT | GTTAAGTCCG | GTTTAGAATA | 3600 |
| :5 | CACAACGCAA | TATAATCAAC | AATTTATCGA | TACATTAAAA | GAGATTGAGA | AGAATGAAAA | 3660 |
| - | TGTTGATTTA | TCAAAAGAAA | TTGATAAGGT | AAAAGCAGCT | AATAATCGAA | TTAATGAATC | 3720 |
| | ATTAAGGTTA | GTTAATCAAT | TAAGCAATGC | ATTAAAGAAT | GGTAGTTCAG | GAACTGCTGA | 3780 |
| o | AGCTACTAAA ' | TTACTAGATC | AACTTTCAAA | ACTAGATTCA | TCATTATCAT | CATTTAGAGA | 3840 |
| | TTATGTTAAA | AAAGATCTTA | ACAGCTCTTT | AGTATCAATA | TCACAACGTA | TTATGGATGA | 3900 |
| | ATTGAACAAA | GGGCAAACTG | CATTATCCAA | TGTTCAGTCT | AAATTAAATA | CAATTGATCA | 3960 |

| | AACAGTATTA | CCAAGTATTG | AACAACAATA | CATTAGTGCT | GTTAAAAATG | CTCAAGCAAA | 4080 |
|----|------------|------------|------------|------------|------------|------------|------|
| _ | CTTCTCGAAA | GTGAAAAGTG | ATGTAGCTAA | AGCTGCTAAC | TTTGTGCGCA | ATGACTTACC | 4140 |
| 5 | ACAGTTAGAA | CAGCGATTAA | CTAATGCGAC | AGCAAGTGTG | AATAAAAATT | TACCAACGTT | 4200 |
| | ATTAAATGGT | TATGATCAAG | CGGTAGGATT | ACTAAATAAA | AATCAGCCAC | AAGCGAAAAA | 4260 |
| 10 | GGCTTTATCA | GATTTAGCTG | ATTTTTCTCA | AAATAAATTG | CCTGATGTTG | AAAAAGATTT | 4320 |
| | GAAAAAAGCG | TTTAAAATTT | TCAAGAAATT | AGACAAAGAT | GATGCAGTCG | ACAAATTAAT | 4380 |
| | CGACACACTT | AAGAATGATT | TGAAAAAGCA | AGCGGGTATT | ATTGCAAATC | CTATTAATAA | 4440 |
| 5 | GAAGACTGTT | GATGTTTTCC | CAGTTAAGGA | TTATGGTTCA | GGTATGACAC | CATTCTATAC | 4500 |
| | TGCACTGTCA | GTATGGGTAG | GTGCACTCTT | GATGGTAAGT | TTATTAACGG | TTGATAATAA | 4560 |
| | ACATAAGAGT | CTAGAGTCAG | TCTTAACGAC | AAGACAAGTG | TTCTTAGGTA | AGGCAGGATT | 4620 |
| 20 | CTTTATAATG | CTTGGTATGT | TGCAAGCACT | CATTGTATCG | GTTGGAGATT | TGTTAATCCT | 4680 |
| | AAAAGCAGGA | GTTGAGTCAC | CTGTATTATT | TGTACTTATA | ACGATTTTCT | GTTCGATTAT | 4740 |
| | TTTCAACTCA | ATCGTATATA | CGTGCGTATC | ATTACTTGGT | AACCCAGGŢA | AAGCCATTGC | 4800 |
| ?5 | AATCGTATTG | CTTGTATTAC | AAATTGCAGG | TGGTGGGGGA | ACATTCCCAA | TTCAAACTAC | 4860 |
| | GCCACAATTT | TTCCAAAACA | TTTCGCCATA | CTTACCATTT | ACGTATGCAA | TTGATTCATT | 4920 |
| 30 | ACGTGAAACA | GTAGGCGGTA | TTGTTCCGGA | AATCCTAATT | ACAAAATTAA | TTATATTAAC | 4980 |
| | GTTATTTGGT | ATAGGATTCT | TCGTTGTAGG | TTTAATTTTA | AAACCTGTAA | CAGATCCATT | 5040 |
| | GATGAAGCGC | GTATCTGAAA | AAGTTGACCA | AAGTAACGTT | ACAGAATAAA | AATTAAATCC | 5100 |
| 15 | ACACATTAGG | GTTATAGCTC | CTTAATGTGT | GGATTTTTAT | GTTTTTAGAC | AGAAGAGATA | 5160 |
| | GTAATTTCTG | TCTTTTATGG | GACGGTTGTT | ATCATTGCTA | TTATCCAGGA | TGACTTACTA | 5220 |
| | TAGGACTAAT | ATTACCGACA | AAGTGAATAT | CCTCGTCTTC | CGTAGTTAAA | ATAAAGCTAG | 5280 |
| 10 | AACCTTTTTG | GATGTCATAG | TGCTTATCGT | TTACTGTTAA | AGTACCAGTA | CCATCGATAA | 5340 |
| | TTGTAACTAA | GCAATAAGCA | TGTGGTTTAT | TGAATTTTAA | ATCTCCATGA | ATATCCCATT | 5400 |
| 15 | TATATACTGC | AAAATATTGA | TTATCTACAA | ATTGAGTTAC | AGTGTGTGTG | TCGATGTGAG | 5460 |
| | TTGTTATAGG | AGTAGTATTT | GGTTCATGAT | TGCCTAATTC | AATCACATCT | TTACTTTGCT | 5520 |
| | CTAAGTGCAA | ATCACGCAAT | TGACCATTTT | GATCTCGTCT | ATCATAGTCA | TAAATACGGT | 5580 |
| 50 | ATGTCGTATC | GGAGGATTGT | TGTGTCTCTA | AAATTAAAAT | ACCCGAACCA | ATGGCATGGA | 5640 |
| | CAGTGCCAGC | AGGAACATAA | TAAAAGTCAC | CGGGCTTAAC | AGGTATACGT | TTGAAAAGAC | 5700 |
| | TGTCAAATTC | ATGATTATCA | ATCATGTCTA | TTAACGTCTG | TTTATTATGT | GCATGTACGC | 5760 |

| | GTTCGCCTTC GTGTTTTAAA GCGTAGTCAT CATCTGGGTG AACTTGAACA GATAATTTAT | 5880 |
|----|---|------|
| _ | CATTGGCATC TAATACTTTA GTTAGCAGAG GGAAACTATC TCGTGAATCA TTATCGAATA | 5940 |
| 5 | ATTCACGATG TTGTGACCAA AGTTGATCTA GGGTCATATC CTTGTATGGA CCATTGATAA | 6000 |
| | TTGTATTAGG ACCATTTGGA TGTGCAGAAA TTGCCCAGCA TTCACCAGTT GTTTCATTAG | 6060 |
| 10 | GGATATCATA GTTAAATGCT TTTAATGCAT GACCGCCCCA AATTCTGTCT TTAAAAACGG | 6120 |
| | GTTGTAAAAA TAATGCCATA GTTAAAACTC CTCTATATTT TCATTAATAA GTTATAAATT | 6180 |
| | TCTGTAGTAC TGTTTGCATT AATTAGTGAT TGGCGTGTCT CATCATTCAT TAACGCTTTA | 6240 |
| 15 | GATAAGCGCT GAAGTATTTT TAAATGTGTA TCCTGACTGT TGTTTGGTAC GGCAATTAAG | 6300 |
| | AATATCAATT GAGGTAGACT ACCATCTAGA CTGTCCCATT TAACACCATG ATTATTTTTC | 6360 |
| | ATAACAGCTA CAATCGGTTG TTTTACAACA TCAGACTTTG CATGTGGAAT GGCCACGTTC | 6420 |
| 20 | ATGCCAATAG CTGTCGTAGm tCcATTTCAC GTTCTAGTAT TGCATTTTTT AAATGCGATG | 6480 |
| | TGTGCTCTAC ATAACGGCAA ATTTTAAGTT TATGAATCAA CATATCAATT GCTTCGTTTC | 6540 |
| | GAGACATGTC GTGATCAGTA ATTATCATAG TTTGTTGATC AAAAACATGA GAAGGTTTAT | 6600 |
| 25 | TGAGATGTGA ATGTTTCGCG GTGTTATCTA CATTGTCAAC CTCTGTATCA TGTTGTGTAA | 6660 |
| | TATCTGTATC ATGAAGTTGC GTGTGTTGCG CTGGTGCATC TACTGCTATA ACTGGTGTAT | 6720 |
| 20 | TGCGTTTTAA TAATAGTACA GTAGTCATTG TGACAAGACT ACCTACTATC ACTGCAAAGA | 6780 |
| 30 | TAAACCATAA TACATGATCA ATACCACCTA ATACAGCCAC GATTGGACCT CCATGTGCGA | 6840 |
| | CTCTATCGCC GACACCACCA ATGGCTGCAA TGACTGATGC AATCATTGCA CCAATGATGT | 6900 |
| 35 | TTGCAGGTAT AATGCGCAAT GGATCTTGGG CTGCGAAAGG AATAGCACCT TCAGTAATAC | 6960 |
| | CAAATAGTCC CATAGTGAAG GAAGCCTTAC CCATTTCTCT TTCGGAATGA TTGAATTTAT | 7020 |
| | ACTITIGAAC AAACGTIGCT AAACCTAAAC CGATIGGIGG TGTACATACA GCAACTGCGA | 7080 |
| 10 | CCATACCCAT AACGGCGTAA TTACCTTCAG CAATAAGTGC TGAGCCAAAT AAAAATGCTA | |
| | CCTTGTTTAC TGGACCGCCC ATATCGAAGG CAATCATCGC ACCTATAATC ATCGCAAGTA | 7200 |
| | TAATAATATT AGCACCTTGC ATACTTTTTA ACCAGGTTGT TAATGCCTCA AAAATATTAG | 7260 |
| 15 | AAATTGGTGC ACCGATTAAA AATATAAATA TCAATCCTAC AACGACCGAT GAAATAATGG | 7320 |
| | GAATAATAAT GATAGGCATA ATTGGTGCCA TTGCTTTTGG AACTTTAATA TCTTTAATCC | 7380 |
| | ACTITGCGAT ATAACCTGCT AAGAAACCAG CAACAATACC ACCTAAAAAT CCTGCGCCTG | 7440 |
| io | CATCACTGCC ATAAAAACTA CCGTCAGCAG CGATAGCGCC GCCAATCATA CCAGGAACAA | 7500 |
| | GACCGGGCTT GTCAGCGATA CTAACAGCGA TATATCCAGC TAGTATTGGA ACCATAAATT | |

| | ATCCTTTTGA | TGTCGTTtCA | CCGCCTAGAG | TCAGCGCGAT | GGCGATAAGG | AGTCCACCAA | 7680 |
|----|------------|-------------------|------------|------------|------------|------------|------|
| | CTACGATAAA | AGGAACCATA | AACGATACAC | CGTTCATTAA | ATGTTGATAC | ACCATTTGAA | 7740 |
| 5 | TACCATTTTT | AGACTTACCG | CGATCTTTCG | AATGATAATT | TGTTTCAGAT | TGATAAATAG | 7800 |
| | GCGCATCTTG | ATTAATGATA | CGTTGAATTA | GACCTCTCGG | ATTATGAATC | CCTTCGCGAA | 7860 |
| 0 | CATTTTCATT | AATCAACCGT | TTACCAACAA | ATCGGGACAG | ATCAACTIGT | TTATCAGCTG | 7920 |
| Ü | CAATTATGAC | ACCGTCAGCT | TCTTCGATGT | CTTGCGTAGT | TAAAACATTT | TCAGCACCAA | 7980 |
| | CACCGCCCTG | TGTCTCTACT | TTAATATCCA | CACCCATTTC | TTTTGCTACC | TGCTCAAGCT | 8040 |
| 5 | TTTCTTGAGC | CATATATGTA | TGTGCAATGC | CATTTGGGCA | TGAGGTAATA | GCTACAATTT | 8100 |
| | TCATAAAATC | ATCTCCTTTT | CTATATTGTA | AGCGTATTCT | CGATACTAAA | AAAAAGAATA | 8160 |
| | ATTACCGTTA | CTAGTGGCAA | TTATTCTTGT | AAGTATTCAA | ATAACTGTTG | CTTTAAACTA | 8220 |
| 20 | TGATCATCTA | AACTACATAA | ATGGTTCACT | GAATCATCAT | CCAAGTTAGC | AATTAATTGC | 8280 |
| | ATCATTTGTT | TTGTAAAAGC | TTTGTCTTTA | TGCGAAATCG | CTAAGAAAAA | GACAAGTTTG | 8340 |
| | ACATCGTGTT | GTCGCCAAGG | AAAAACATCT | TTTGTGCGAA | AAATAAGCAC | ATGTGATTGT | 8400 |
| ?5 | AAAACTTTTT | CAGGATCTCC | ATGAGGAATC | GCCATAAAAT | TACCTATGTA | TGTAGAAGAT | 8460 |
| | GATTTCTCAC | GCTCTAAAGC | TGATTCGATA | TATCCTTCTA | CAATCGCATG | ATGTGCTTGT | 8520 |
| 30 | AATATTTTTT | GAGCTTCTTC | AAAAATTTGC | ACAGTATGCC | GTGATTTTTG | TTCAGTATTT | 8580 |
| 30 | ACGACAAGGA | AATTGACAGT | GTCCATATGA | TGATGTGCTT | GAACCGGATT | TTGCTTTTGC | 8640 |
| | TTCACAACGT | GTCTGATTTT | GTGACGATCA | TCTTCAGAAA | ATAATGGTGC | AACCTTGATA | 8700 |
| 35 | GTCGTCAGGT | GCTTAGGAAG | TATGTTTAGC | GTTTGTTTAG | GAATATCATG | GGTCGTTATT | 8760 |
| | AATAAATCTA | CATTGTCAAA | GTGATAGTGT | GTTATATTTT | CTAGTTTAAT | CGTATTTATC | 8820 |
| | ACTGÁCAACT | CȚTCGGATAA | GTTATTTATT | TTAGTTTCTA | AAAAATTCGA | CACACCTAGA | 8880 |
| 40 | CCATAATAAC | AAGCAATGAC | TACATTTAAT | TGTGTTTTGG | TACGACGCTC | GATGGCAGCT | 8940 |
| | TGAAAATGAA | TTGTTAAAAA | TGCAATTTCA | TCTTCGCTCA | TCTCTATATC | AGTATCAATT | 9000 |
| | GCTAATTTAT | CAATCGCTTC | AAAAAGTGTG | TTAAACACAA | AGGGATAGAG | TTTTTTAATC | 9060 |
| 45 | TCTATAACTA | AAGGATTGTT | TAAATAAATG | TTTTGAGTGA | TACGTAAATA | TGCTTTACTA | 9120 |
| | AAATGATTAT | ATAAATTTTG | TTGTAAAATC | GAATCTTCAT | TGAAAGGTAC | ATGAATACGT | 9180 |
| | TGCTGCATCA | ATTCGATTAA | GCGATCAATA | TAACTTTGTA | TAAATATACG | TTCTATGCCA | 9240 |
| 50 | ATATCGAGTT | TATTAAAATG | ATAAGCAATA | AAGAATGAAA | ACATATTGAT | TACTTTTTCG | 930 |
| | TTCAAGTCAT | ДАССТААТСТ | TTCGTTGATT | TGCTTAATGC | AAGATTGAGA | TATCAATTTT | 9360 |

| | - | |
|----|--|------|
| | AGATGAATTA AAAGCTGTTG TATTTGAATA TCAGTTGTTT CAATACTATG TTGTTGAAGT | 9480 |
| 5 | GTCTCTTGTA TAATATGCGA AATCATCCTT TGGTGTGAAT CAGGTAATTC aTTTAAAATT | 9540 |
| | AGGTCTTCAA CATGTACATG CCCTGATGAT AATTGATTTA AATGGATGAT GGCATTAGTG | 9600 |
| | ATATCATTAT CTGTTCCATC GAC | 9623 |
| 10 | (2) INFORMATION FOR SEQ ID NO: 167: | |
| 15 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1021 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 167: | |
| 20 | ACCGTGGAAA CACGTCTAGT CAATCAGAAA GCGATAAAAA TGTGACTAAA TCATCTCAAG | 60 |
| | AGGAAAATCA AGCAAAAGAA GAATTACAAA GCGTTTTAAA CAAAATTAAC AAACAATCAA | 120 |
| | GTAAGAATAA TTAAAAAATT TTGATATTGT CTATGTTTAT AGTTCACAAG CCATTCAACG | 180 |
| 25 | TATTGTAAAC TAAGGATAGT GTATTTTTT AATAGTAATT TGTCAGGAGG TGCCTATCTA | 240 |
| | TGGAAGAACA TTACTACGTA AGTATTGATA TTGGATCATC AAGCGTAAAA ACAATAGTAG | 300 |
| 30 | GCGAGAAATT TCACAATGGT ATAAATGTGA TAGGTACAGG ACAAACCTAC ACGAGCGGTA | 360 |
| | TAAAAAATGG TTTAATTGAT GATTTTGATA TTGCGCGACA AGCAATCAAA GACACAATTA | 420 |
| | AAAAGGCATC AATCGCTTCG GGTGTTGATA TTAAAGAAGT TTTCCTGAAA TTACCTATCA | 480 |
| 35 | TTGGAACGGA AGTTTATGAT GAATCAAATG AAATCGACTT TTATGAGGAT ACAGAAATCA | 540 |
| | ACGGTTCACA TATCGAAAAA GTATTAGAAG GTATTAGAGA AAAAAATGAT GTGCAAGAAA | 600 |
| | CAGAAGTAAT TAATGTGTTC CCGATTCGTT TTATAGTCGA TAAAGAAAAT GAGGTTTCAG | 660 |
| 10 | ACCCTAAAGA ATTAATTGCC AGACATTCAT TAAAGGTTGA AGCAGGCGTA ATTGCTATTC | 720- |
| | AAAAATCGAT TTTAATTAAT ATGATTAAAT GCGTAGAAGC ATGTGGTGTT GATGTATTAG | 780 |
| | ATGTTTACTC TGATGCATAT AACTATGGTT CAATCCTAAC AGCTACTGAA AAAGAGTTAG | 840 |
| 15 | GTGCATGTGT CATTGATATT GGTGAAGACG TTACGCAAGT TGCTTTTTAT GAACGCGGTG | 900 |
| | AATTAGTAGA TGCTGATTCT ATCGAAATGG CAGGGCGTGA TATTACAGAC GATATTGCAC | 960 |
| :0 | AAGGITTAAA CACTICINAT GAAACTGCTG NAAAAAGTTA AACACCAATN TGGTCATGCA | 1020 |
| 50 | T | 1021 |

(2) INFORMATION FOR SEQ ID NO: 168:

(A) LENGTH: 7963 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 168:

| 10 | TAATCTATTA | TAAAAACTGT | CCATACCCTT | TGATTACCTT | CTCTTCAGGT | ACAGGCCACA | 60 |
|----|-------------|------------|------------|------------|------------|------------|------|
| | CTTGAGGCCA | TAAGCCATAT | GCTTGCTGTG | AATAAAATTG | TGCCATTTGT | AACAATATAA | 120 |
| | TATATACAAA | TAAACACCCA | ATAATTGCTG | TCACTAATGG | ATATGATAAC | CAAACCATTA | 180 |
| 15 | ATAAAACTGC | AATAATTACT | AACCTAAAGA | TAATATTAAA | TGCGTCTCTC | CCTCTTATAA | 240 |
| | AGCTTCTAAT | AAATAAGAAT | AAATACATCG | CATTAGAGTT | AAATTTACTA | CCCTTTGGAA | 300 |
| | CTGGTAAAAG | TATATCTAGA | TAACTTCTTC | TGACTGCAGA | TTCTTTCAAA | TGTTTTACAT | 360 |
| 20 | CGGTGAACAT | ATTAACAAAT | TTATAATAAT | TCATATGATG | TCGATGTTCG | ATTGCAATCA | 420 |
| | TTTTCTCCCA | AGGATACAAA | AAGCCTGGTT | TATATTTTTT | AACTAAAAAT | TCTATTAACA | 480 |
| 25 | CAGGCAAAGC | AACCATCACA | AATGCGATGT | ACCATTTTGG | AGCTAATAGT | AAGTAATATG | 540 |
| 20 | TTAGAGCAAA | GGTGATGAAT | GATATTAAAT | TAACTTGCCA | TGTTTTAAGT | CCCGATTGAT | 600 |
| | ACCATTGCCA | TCTTAAGCGT | AAACCAACAT | ATGGAAAAAT | TAATGCACTG | ACTCCAAAAC | 660 |
| 30 | AAATATAAAA | TGCCACATTA | TGTTGATTAA | TATTGTAAAA | CAACGGGAAC | ATTACAATAA | 720 |
| | CAATAATGAG | TTGGATTAAT | ATGCGCGCAA | AGTAACTATA | TAAAATCGCA | TGACGCATAA | 780 |
| | ATTGAGACAT | GTGTTTTTCA | AATGGTAATA | AAAAGATTTT | ATCCGCTTCT | TTTAACAGTG | 840 |
| 35 | GTCsCmTTGG | AAAAATAGrT | GTCAACGCAA | CAATCACTGC | TGCTATTaAT | GAAAAATTGa | 900 |
| | TATTCGTTGG | AATATGTTTT | AACCATTCAC | CATATCCArA | AATAAATGCA | CCCAGCAAAA | 960 |
| | TAAGTAAAAA | GACCATGAAA | TGACCATTAA | ATATAAACTT | ATTATAATAA | TTTTtCTCTT | 1020 |
| 40 | TACGAAGGGC | ATGTAATCTT | TTATTAAATA | ATGTGGTAGC | TTGGTTACGC | ATGTACATCT | 1080 |
| | CCACCTTGCG | TCACATGAAT | ATATATATCG | TCTAATGTTT | GATTATGTAA | GCCAGTTTGT | 1140 |
| 46 | TGTCTCAATG | CTTCTAAATC | TCCAAATGCA | ACGACTTCAC | CTTCGTCTAG | TATGATAAAA | 1200 |
| 45 | CGATCACAGT | AACGTTCAGC | TGTTGCTAAA | ATATGTGTAC | TCATTAGAAC | GGTTCTACCT | 1260 |
| | TCGTTTTTCT | TTTCAACCAT | TAAATCTAAC | ATGGATTGAA | TTCCTAATGG | ATCTAGGCCA | 1320 |
| 50 | AGGAATGGTT | CGTCTATAAT | ATACAATTCG | GGATTAACGA | TAAACGCACA | AATAATCATG | 1380 |
| | ACTITITIGIT | TCATCCCCTT | AGAAAAATGA | CTCGGAAAAA | CTTTCAACTC | ATTTTCTAAA | 1440 |
| | CGGAATGTCT | TTAATAATGG | CATTGCTCGA | TTCATCGTTT | CATCACGATC | AATATCATAT | 1500 |

55

| | TCCGGAATAT | ' AAGATAACTT | TCTTCTATAL | GCCTCTATG1 | CATCATTAAT | GTTGATATCT | 1620 |
|------------|--------------|--------------|------------|------------|------------|------------|------|
| | GAAATTGATA | GAGATCCTTC | CATAGGTGTA | AGCAATCCTA | GCATATGTTT | AATCGTTGTA | 1680 |
| 5 | CTCTTACCAG | CGCCATTAAG | GCCAATAAGI | CCAACAATTT | CGCCTTTGTT | TAATTCAAAA | 1740 |
| | TTTATATCTT | TAATTACAGG | GCGTTTTCCA | TATCCACCTG | TAAGCTGTTC | TACTTTAACT | 1800 |
| 10 | GTCATAAGGC | ACCTCCATGA | CTTATATTGT | ACCAAAAATT | ATAAAATGÇT | CATATTAAAT | 1860 |
| | ACACATGTCC | TAATATCGAA | TTTTTAGCGA | CAATGTTATA | ATGAATGGTA | ATACTAGTTG | 1920 |
| | AAAAGGAGTG | TAGTCATCAT | GTCAGAAACA | ATTTTCGGCA | AAATTTTAAC | TGGAGAAATT | 1980 |
| 15 | CCTAGCTTTA | AAGTATATGA | AGACGATTAT | GTCTATGCCT | TTTTAGATAT | ATCACAAGTT | 2040 |
| | ACTAAAGGAC | ATACGTTATT | AATTCCTAAA | AAAGCTTCTG | CTAATATCTT | TGAAACTGAT | 2100 |
| | GAAGAAACAA | TGAAACATAT | CGGTGCAGCA | TTACCTAAAG | TAGCAAATGC | TATTAAGCGT | 2160 |
| 20 | GCATTTAATC | CTGATGGTTT | AAACATTATT | CAAAATAATG | GTGAGTTTGC | AGATCAATCT | 2220 |
| | GTATTTCATA | TTCATTTCCA | CTTAATTCCT | CGATACGAAA | ATGATATTGA | TGGATTTGGT | 2280 |
| 25 | TATAAGTGGG | AAACACATGA | AGACATTTTA | GATAACGATG | CAAAACAACA | AATTGCTGAA | 2340 |
| | CAAATTCAAG | CACAATTTTA | AATGTATGCT | TAATCTAAGC | TCGAACGGGT | ATAATATGAT | 2400 |
| | TAATATTATA | ACAATTGCGT | TTGAAGTGAT | AACATCAAGG | TTAGCAATTT | TAAACAAAAT | 2460 |
| 30 | GAGTTATCAA | GATAACAGAT | GTTAAAAGTG | AGGAGAATAT | AAATGAAAGC | ATCACGCATT | 2520 |
| | CTATTCGGTA | TCGGTGTTGG | CGTAGCAGCT | GGTTTTGTAG | TTGCACTTCA | AGGACGTGAC | 2580 |
| | GACAAAAGTG | TCAAGAACAA | CACGATCGAT | CGTACTGCCC | CTACTGGTTC | AAAATCAGAA | 2640 |
| 35 | CTACAACGTG | AATTTGAAAC | GATTAAACAA | AGTTTTAATG | ACATTTTAAA | CTATGGTGTT | 2700 |
| | CAAATTAAAA | ACGAAAGTGC | GGAATTTGGT | AGTTCAATTG | GTGGTGAAAT | TAAGTCATTA | 2760 |
| | CTTGGAAACT | TCAAATCTGA | CATTAATCCT | AATATTGAAC | GTTTACAGTC | ACACATCGAA | 2820 |
| 10 | AATTTACAAA | ATCGTGGCGA | GGATATTGGA | AACGAAATTT | CTAAGTAGCA | GGTTACGTTC | 2880 |
| | TCGATCACAA | CTATTTTTAT | TAGTAACAGC | ATATTTATTT | ATTAAAATTA | AATGCCAAAT | 2940 |
| . 5 | AAACGAGATG | ACATTAGAAA | TTAGATATTT | CTTGTCATCT | CTTTTTTAAA | ACTCAAATGA | 3000 |
| | ACTTATGTTT | ACAAATTATA | GGAAGACATT | GTTTGTAGTG | ATTTTCGCTT | AAATCATATT | 3060 |
| | TATGAATTGA | TTGAAAACAT | TGCTTAGGAT | TCATTGTGTT | ATCCLTGCAC | TTTGATTACG | 3120 |
| | CTTTACTTAA | ATCATTATCG | ACAAACAACA | TACTTATATT | TTCATTGAGC | CGAACCTTAT | 3180 |
| 0 | ATACACATTA (| CATATACCTT : | ACTTGCACAA | ATTATTAATC | TGGTGTTTAT | TATAATTACA | 3240 |
| | TATCACTATA | TTTTTAGCAT | TTGTATAACT | TAGTTGGTCA | AAAGATGCTT | TTGCATATGC | 3300 |

| | TTTCATAAGT | GATGCTTTAT | TAGCAAGAAT | ATGTGTTCGC | AGAAATTTGT | TCTGCATTCT | 3420 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | ACTTCTACGC | TAGTCAATCA | GACAATTTTA | CCAATCCCCA | CTTTCGCGTT | TCAAATCAAA | 3480 |
| 5 | CAATACGTCG | CTCCTTTCTT | CTTATATAAC | AATTCTTCTA | ACATGATATG | TTACTATTGA | 3540 |
| | ATTACTGAAC | CTGAGTTAGT | TATAATCTAA | CTTATATTGA | AAAGAGATGA | GGCGTAAGAT | 3600 |
| | ATGTTTTTAT | GTAAAAGACA | AATTGATATC | AATGCACGAT | TTGGTTTGCC | TAGAATTGCA | 3660 |
| 0 | TTTATGAGTG | CAGTTGCAAC | CATCATTATG | TTTTTAGTTA | GTTATGAAGT | AATGTATTTT | 3720 |
| | TTATCTAATA | CGCCATTATC | AGATAGACAT | TTTCTCATCT | TTTTATTACT | TGTATTTATG | 3780 |
| | ACGTATCCAT | TACATAAAAG | TATACATTTA | TTATTTTTCT | TACCATATAG | AAAATCGTTT | 3840 |
| 15 | AAAGTTCATA | AGTTAACTAA | AAGAAAATGG | CTTATATTCT | ATAATACCTA | CGTCAATCAA | 3900 |
| | CCTGTACACA | AATTTTATTT | TTGCATTAAC | TTAATATTGC | CGTTAATTAT | CTTATCTGCA | 3960 |
| 20 | ATGTTCGTTT | ATCTAACAAT | TTCATTCCCG | CAATATGGAC | ATTATTTAT | GTTCTTATTG | 4020 |
| | GCATTGAATT | TCGGTATTTC | CATTACAGAT | TTATTATATT | TAAAAATAAT | TATATTTTCT | 4080 |
| | AATTATGGAC | aatatataga | AGAACATAGT | ACAGGTATTA | ATATTTTGAA | AAAAATTAAA | 4140 |
| ?5 | AATCCATATC | ATTTATAACA | AAATAATTAT | AGCAAGGTGT | TATTATTTGT | TTTTAGGCTA | 4200 |
| | TGTAATAgcT | tACAATCAAA | TGTATATAGA | CCTTGTTTTT | TTATTTTCAT | CAATTTCTAC | 4260 |
| | CCCTAAACCT | AATGCTCTAG | TCTGATGTCA | TGGGTTATTG | ATTGGTGATA | АТАТААААСТ | 4320 |
| 30 | ATGTTATATT | CACGATGATT | AACTTACAAA | GGAGTTTCAA | CTATGAAGAT | GATAAACAAA | 4380 |
| | TTAATCGTTC | CGGTAACAGC | TAGTGCTTTA | TTATTAGGCG | CTTGTGGCgC | TAGTGCCACA | 4440 |
| | GACTCTAAAG | AAAATACATT | AATTTCTTCT | AAAGCTGGAG | ACGTAACAGT | TGCAGATACA | 4500 |
| 35 | ATGAAAAAA | TCGGTAAAGA | TCAAATTGCA | AATGCATCAT | TTACTGAAAT | GTTAAATAAA | 4560 |
| | ATTTTAGCTG | АТАААТАТАА | AAATAAAGTT | aatgataaga | AGATTGACGA | ACAAATTGAA | 4620 |
| 10 | AAAATGCAAA | AGCAATACGG | CGGTAAAGAT | AAATTTGAAA | AGGCCCTTCA | ACAGCAAGGT | 4680 |
| •0 | TTAACAGCCG | ATAAATATAA | AGAAAATTTA | CGTACTGCTG | CTTATCATAA | AGAATTACTA | 4740 |
| | TCAGATAAAA | TTAAAATCTC | TGATTCTGAA | ATTAAAGAAG | ACAGCATGAA | AGCTTCACAC | 4800 |
| 1 5 | ATTTTAATTA | AAGTTAAATC | TAAGAAAAGC | GACMAAGAAG | GCTTAGATGA | TAAAGAAGCG | 4860 |
| | AAACAAAAAG | CTGAAGAAAT | TCAAAAAGAA | GTTTCAAAAG | ATCCAAGTAA | ATTTGGTGAA | 4920 |
| | ATCGCTAAAA | AAGAATCAAT | GGATACTGGT | TCAGCTAAAA | AAGATGGCGA | ATTAGGTTAT | 4980 |
| 50 | GTTCTTAAAG | GACAAACTGA | TAAAGATTTT | GAAAAAGCAC | TATTTAAGCT | TAAAGATGGT | 5040 |
| | GAAGTATCAG | AGGTTGTTAA | ATCAAGCTTT | GGATATCATA | TTATTAAAGC | TGATAAACCA | 5100 |

| | AAAAATCCAA | AATTATTGAC | TGATGCATAC | AAAGATCTAT | TAAAAGAATA | CGATGTTGAC | 5220 |
|-----------|-------------|-------------|-----------------|------------|------------|------------|------|
| | TTTAAAGATC | GTGATATTAA | ATCAGTTGTC | GAAGATAAAA | TCTTAAACCC | TGAAAAACTT | 5280 |
| 5 | AAACAAGGTG | GCGCACAAGG | CGGACAATCC | GGCATGAGCC | AATAACACAA | AACCGAGCGA | 5340 |
| | CCGTGGTTCA | AAAATCATAC | CACGGCCGCT | CGGTTTTTTC | GCATTAAAAA | TCGGACAGAT | 5400 |
| | GAGCTCATGT | TTCAGTATAC | TCATCTGTCC | GATATCTTTT | AATTCTTAAT | CGAGTGATTC | 5460 |
| 10 | AGGATTGTAG | AATCTACGAT | TTTCAAGACC | AAATATTTTA | TCTGTAAACT | GACCCTTGTC | 5520 |
| | AGTTTTTTTA | TATGCCTTTT | CAAACATATT | CATTCTAGCA | TCGATATTAT | CGATATAGCA | 5580 |
| 15 | TAAAATTTCT | GCTTCTTTTA | AGTATGGCAG | TTTTGGAGAA | CCATACTCTA | ACTTACCATG | 5640 |
| | ATGAGATAAA | ATCATATGTC | TTAACAACAT | GATTTCTTCT | CCTTCAATGT | TCAATTCACG | 5700 |
| | AGCTGCTTCA | ACTACTTCAT | CACTCGCAAT | CGAGATGTGT | CCTAATAAGT | TACCTTCGAC | 5760 |
| 20 | TGTATACGAC | GTCGCAACAG | GACCACTCAA | TTCTCTAACT | TTACCAATAT | CATGCAAAAT | 5820 |
| | AATACCACTA | TATAACAAAC | TTTTGTTTAA | CAATGGATAA | ATGTCaCAAA | TTGATTTTGC | 5880 |
| | AATACGTAAC | ATCGTTAATA | CATGATAGCT | TAAGCCACTC | GCAAAGTTAT | GaTGATGAGA | 5940 |
| 25 | ACTAGCAGCT | GGATATGTGT | AAAATCGTTC | TTGATATTTT | TTCAATAAAT | GACGTGTGAT | 6000 |
| | ACGTTGTAAA | TTAGCATTTT | CAATATCTAG | CAAATAATGA | GAAATCTCTT | CTTGTATTTC | 6060 |
| | TGCCGGTGAT | AAAGGTGCAC | CATCTACAAA | TTGTTCTGTT | TTTAATTGAT | CTTCAGTTGT | 6120 |
| 30 | CGCTAGTCTA | ATTTGGTTGA | CTTTCATCTG | TTTATTTCCG | CGATAGTTTA | TGATGTCACC | 6180 |
| | TTTAACATGT | ACAATTTCTT | CAGGCTTGAT | TGTTGCCATA | TCATTTTTTG | TAGCCGTCCA | 6240 |
| <i>35</i> | AAATTTCGCT | TCAATTTCAC | CACTITTATC | TTGCAAATGT | AATGTCATAT | AATCTTTACC | 6300 |
| | TTGTGCTGTT | ACACCCTGTG | TAGCTTTATG | CACTAAGAAA | AAGTGATCAA | CTGAATCTCC | 6360 |
| , | GGGATTTAGA | TTCTCTATAT | TTCTCATCGT | TTCCCGCCTT | CCTCTATTTT | GTTTAATGTA | 6420 |
| 40 | ATCACTTCTT | TTGATGGAAC | AATATTATCT | TTTACACATG | TAAAGTATAG | TACTTGATAG | 6480 |
| | TGTTCTGATA | ATGATCGTAA | ATAATTCAAC | ATTTTTCAG | TACGTTTTTT | ATCAAAATGA | 6540 |
| | ACAAATGCAT | CATCAACAAT | TAATGGGAAC | GGATAATATG | GTCTTAGTAC | СТТААТТААА | 6600 |
| 45 | CTGATACGTA | AAGCTACATA | AAGTAATTCT | TTTGTAGATT | GACTTAGTTC | AACAGGATCA | 6660 |
| | TATAATTGAC | CATTAACATG | TTTAACCGTA | ATTGAATCTT | CATTATAGTT | AATCATCGTA | 6720 |
| | TATCTGCCAT | CTGTTAAATG | CTTCAATATT | TCTACCGCTT | CATTAATAAC | TTGAGGCAAA | 6780 |
| 50 | CGTTTATCTT | TAATTTGTTT | AATGTGTTCA | TCAACTAAAC | TTTGTAAATA | ACTTAAACTT | 6840 |
| | CCCCA ATCTT | TTCCCARDADO | a mma a commo a | | | | |

| | GCTTGCATTT | CAAGATATTG | CTCATTATAT | TCGTCAACTT | GAGTAGCCAA | TAAATGATCT | 7020 |
|----|------------|-------------|--------------|------------|------------|------------|------|
| | TCTTCTTCAA | GTTGTGCAGT | TGTTTTTCA | CTTAAACTAG | AACTTAATTC | ATAAGAATAG | 7080 |
| 5 | TTTTGGTTCT | CAAGATATTT | AGTTAAATCA | TTAAAACGAC | TCAAATTACT | AGTATAAGTT | 7140 |
| | TGGTAATCTT | CATGATGTTG | GTAAAAATCT | TCTTCAGTAC | CAACATTGAT | AAAATCGAAT | 7200 |
| | AGTGCTGTAA | TTTCTTTATT | ATTTTCTTCT | AATTGAGCAT | TTAAATGATT | TAATTCATTT | 7260 |
| 10 | GTAACAAGTT | TGGTATTTTC | AGCATTAATA | CGCCATTTTT | CATTCGTGTC | TTCAGCTGAT | 7320 |
| | TTCAACCATT | GTtGCACATC | GTGGAATAAA | GATAATTTGT | TGAAATAAAC | AAATTGTGAT | 7380 |
| 15 | TTTGTAACAG | CTTCAGCATG | ATTGTAGAAT | GTATCTAATT | CTTGAACCAA | TTGCTGGCGT | 7440 |
| 15 | TGTTGATTTA | AATCACTGAT | ATGTTGATCT | AATGCTTTAA | TATTCGCCAT | TGTAGAAATA | 7500 |
| | CTATCAACAA | TTAAATCATT | TGAAATTTTA | GATGATAAGT | ATAATTCATC | CTTAACGTTC | 7560 |
| 20 | TCAACTGTCG | ATTGTAATTC | ATCATGACGC | CCTTTCGCAT | CATTTAAACG | ACCTTCAATA | 7620 |
| | TACTGACGTT | TCTCTTCTAA | AATATCTTTA | TTTTTCAAAG | CTTGTTGCCA | GTGATCACGA | 7680 |
| | ATGCGATATT | GCTCATCAAG | ATCAAAATCT | AAGTCATAAT | TTTCATCTAA | AATGGCTAGT | 7740 |
| 25 | TGTGCTTTAA | TTTCTTCGAT | TTCATCTGTG | ATGGCCTCGC | TATAATCTAC | TTCTTTTGAT | 7800 |
| | TTAGACATGA | TGATACCGAT | AACAAATACT | AAAGTTAATA | CTGCGAAAAT | AATACCAAAC | 7860 |
| | AACATGTTGT | TTGAAATAAA | TGAGAAGGCA | GTTAAACCAA | TACCTACTAA | TGTTAAAAGr | 7920 |
| 30 | ATAAACGTTG | TTCGKAACAA | TTTTTGACGT | TTTTGttTTT | CTT | | 7963 |
| • | (2) INFORM | ATION FOR S | EO ED NO. 14 | εq. | | | |

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3958 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

40

45

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 169:

ATATTGTCTT TACAATAGTT TGCTATGGAG GTAATTAACC AATAGGAGGA ATTTATAATG 60
GCAGTAATTT CAATGAAACA ATTACTAGAA GCGGGTGTTC MCLLCGGTCA CCAAACACGT 120
CGTTGGAACC CAAAAATGAA AAAATATATC TTCACTGAGA GAAATGGTAT TTATATCATC 180
GACTTACAAA AAACAGTGAA AAAAGTAGAC GAGGCATACA ACTTCTTGAA ACAAGTTTCA 240
GAAGATGGTG GACAAGTCTT ATTCGTAGGA NCTAAAAAAC AAGCACAAGA ATCAGTTAAA 300
TCTGAAGCAG AACGTGCTGG TCAATTCTAC ATTAACCAAA GATGGTTAGG TGGATTATTA 360

| | 0110100000 no estado | - |
|------------|---|------|
| | GAAGATGGTT TATTCGAAGT ATTACCTAAA AAAGAAGTAG TAGAACTTAA AAAAGAATAC | 480 |
| _ | GACCGTTTAA TCAAATTCTT AGGCGGAATT CGTGATATGA AATCAATGCC TCAAGCATTA | 540 |
| 5 | TTCGTAGTTG ACCCACGTAA AGAGCGTAAT GCAATTGCTG AAGCTCGTAA ATTAAATATT | 600 |
| | CCTATCGTAG GTATCGTTGA CACTAACTGT GATCCTGACG AAATTGACTA CGTTATCCCA | 660 |
| 10 | GCAAACGACG ATGCTATCCG TGCGGTTAAA TTATTAACTG CTAAAATGGC AGATGCAATC | 720 |
| ,, | TTAGAAGGTC AACAAGGCGT TTCTAATGAA GAAGTAGCTG CAGAACAAAA CATCGATTTA | 780 |
| | GATGAAAAG AAAAATCAGA AGAAACAGAA GCAACTGAAG AATAATCAAC TGTTGAATCT | 840 |
| 15 | GACTTAGATA TAGTTTAAAT GGGTGATAAG ATATTAATGC TTATCACCTT TTTTAAAAAG | 900 |
| • | AAAATCGAGG CAAATTACAA ATATTCAATT AGAGTATTGG CAATCTTGCC TATAATAATG | 960 |
| | CTAAAATCAT AATATAAAN ATGATAACTT ATTGGAGGAA TAATGAATGG CAACTATTTC | 1020 |
| 20 | AGCAAAACTT GTTAAAGAAT TACGTGAAAA AACTGGCGCG GGTATGATGG ATTGTAAAAA | 1080 |
| | AGCGCTAACT GAAACTGATG GTGACATCGA TAAAGCGATT GACTACCTAC GTGAAAAAGG | 1140 |
| | TATTGCTAAA GCAGCTAAAA AAGCAGACCG TATTGCGGCT GAAGGTTTAG TACATGTAGA | 1200 |
| 25 | AACTAAAGGT AACGACGCAt TATCGTTGAA ATCAACTCTG AAACAGACTT TGTTGCTCGT | 1260 |
| | AACGAAGGTT TCCAAGAGTT AGTTAAAGAA ATCGCTAATC AAGTATTAGA TACAAAAGCT | 1320 |
| | GAAACTGTTG AAGCTTTAAT GGAAACAACT TTACCAAATG GTAAATCAGT TGATGAAAGA | 1380 |
| 3 0 | ATTAAAGAAG CAATTTCAAC AATCGGTGAA AAATTAAGTG TTCGTCGTTT TGCTATCAGA | 1440 |
| | ACTAAAACTG ATAACGATGC TTTCGGCGCT TACTTACACA TGGGTGGACG CATTGGTGTA | 1500 |
| ac. | TTAACAGTTG TTGAAGGTTC AACTGACGAA GAAGCAGCAA GAGACGTTGC TATGCATATC | 1560 |
| 35 | GCTGCAATCA ACCCTAAATA TGTTTCTTCT GAACAAGTTA GCGAAGAAGA AATCAACCAC | 1620 |
| | GAAAGAGAAG TTTTAAAACA ACAAGCATTA AATGAAGGTA AACCAGAAAA CATCGTTGAA | 1680 |
| 10 | AAAATGGTGG AAGGACGTTT ACGTAAATAC TTACAAGAAA TTTGTGCTGT AGATCAAGmT | 1740 |
| | TCGTTAAAAA CCCTGATGTA ACAGTTGAAG CTTTCTTAAA AACAAAAGGT GGAAAACTTG | 1800 |
| | TTGACTTCGT ACGCTATGAA GTAGGCGAAG GTATGGAAAA ACGCGAAGAA AACTTTGCGG | 1860 |
| 15 | ATGAAGTTAA AGGACAAATG AAATAATCTG TCATAAAGTA AAACAAGGAA GAAGACACCT | 1920 |
| | TTAATGTTGC TTTATTAAAA TGTAAATCAT TCTAATAAAA CGACAACTGT GTCTTCTTTA | 1980 |
| | CTTGTATATG TTACATATAT TCACGATAGA GAGGATAAGA AAATGGCTCA AATTTCTAAA | 2040 |
| 50 | TATAAACGTG TAGTTTTGAA ACTAAGTGGT GAAGCGTTAG CTGGAGAAAA AGGATTTGGC | 2100 |
| | ATAAATCCAG TAATTATTAA AAGTGTTGCT GAGCAAGTGG CTGAAGTTGC TAAAATGGAC | 2160 |

| | TTAGGTATGG | ACCGTGGAAC | TGCTGATTAC | ATGGGTATGC | TTGCAACTGT | AATGAATGCC | 2280 |
|------------|------------|------------------|------------|---------------|------------|------------|------|
| | TTAGCATTAC | AAGATAGTTT | AGAACAATTG | GATTGTGATA | CACGAGTATT | AACATCTATT | 2340 |
| 5 | GAAATGAAGC | AAGTGGCTGA | ACCTTATATT | CGTCGTCGTG | CAATTAGACA | CTTAGAAAAG | 2400 |
| | AAACGCGTAG | TTATTTTTGC | TGCAGGTATT | GGAAACCCAT | ACTTCTCTAC | AGATACTACA | 2460 |
| | GCGGCATTAC | GTGCTGCAGA | AGTTGAAGCA | GATGTTATTT | TAATGGGCAA | AAATAATGTA | 2520 |
| 10 | GATGGTGTAT | ATTCTGCAGA | TCCTAAAGTA | AACAAAGATG | CGGTAAAATA | TGAACATTTA | 2580 |
| | ACGCATATTC | AAATGCTTCA | AGAAGGTTTA | CAAGTAATGG | ATTCAACAGC | ATCCTCATTC | 2640 |
| 15 | TGTATGGATA | ATAACATTCC | GTTAACTGTT | TTCTCTATTA | TGGAAGAAGG | AAATATTAAA | 2700 |
| 3 | CGTGCTGTTA | TGGGTGAAAA | GATAGGTACG | TTAATTACAA | AATAAATTTA | GAGGTGTAAA | 2760 |
| | ATAATGAGTG | ACATTATTAA | TGAAACTAAA | TCAAGAATGC | AAAAATCAAT | CGAAAGCTTA | 2820 |
| 20 | TCACGTGAAT | TAGCTAACAT | CAGTGCAGGA | AGAGCTAATT | CAAATTTATT | AAACGGCGTA | 2880 |
| | ACAGTTGATT | ACTATGGTGC | ACCAACACCT | GTACAACAAT | TĄGCAAGCAT | CAATGTTCCA | 2940 |
| | GAAGCACGTT | TACTTGTTAT | TTCTCCATAC | GACAAAACTT | CTGTAGCTGA | CATCGAAAAA | 3000 |
| ?5 | GCGATAATAG | CAGCTAACTT | AGGTGTTAAC | CCAACAAGTG | ATGGTGAAGT | GATACGTATT | 3060 |
| | GCTGTACCTG | CCTTAACAGA | AGAACGTAGA | AAAGAGCGCG | TTAAAGATGT | TAAGAAATT | 3120 |
| | GGTGAAGAAG | CTAAAGTATC | TGTTCGAAAT | ATTCGTCGTG | ATATGAATGA | TCAGTTGAAA | 3180 |
| 30 | AAAGATGAAA | AAAATGGCGA | CATTACTGAA | GATGAGTTGA | GAAGTGGCAC | TGAAGATGTT | 3240 |
| | CAGAAAGCAA | CAGACAATTC | AATAAAAGAA | ATTGATCAAA | TGATTGCTGA | TAAAGAAAAA | 3300 |
| | GATATTATGT | CAGTATAAAA | CTAATATACA | ATGACATATT | AAAATGCCAG | TATTAAACGA | 3360 |
| 35 | TAATGTAACA | TTTAAAATGG | GCATGTTTAA | TTAAATCAAA | GATGCATGTG | AAATTTAATA | 3420 |
| | TTCAGAATGA | GCATAAAAAT | GGTGTTTAAA | CAAGTTAATT | AAACATATAC | TTTATAAATA | 3480 |
| ‡ 0 | ATAGGCATTA | GGTATATTGC | TATAATAAAG | TTATGTAATT | TTTAACCTCA | GTATGTATGT | 3540 |
| | CACATTTCTG | GTGTAAACTG | TACCGAGTCA | GACTTTGGTA | CAGTTTTTTT | ATTTGCTTAT | 3600 |
| | TCAATGCATT | AAATGAGTAT | GATAAAATGA | TAATGATTGT | TTAGTAACTT | ATACTATATG | 3660 |
| 1 5 | ACAGAGATGA | TCAGGCTCGG | AGGAAAGACC | ATGTTTAAAA | AGCTAATAAA | TAAAAAGAAC | 3720 |
| | ACTATAAATA | ATTATAATGA | AGAATTAGAC | TCGTCTAATA | TACCTGAACA | TATCGCTATT | 3780 |
| | ATTATGGATG | GTAATGGGCG | ATGGGCTAAG | AAGCGAAAAA | TGCCTAGAAT | TAAAGGTCAT | 3840 |
| 50 | TACGAAGLAT | GCAAACAATA | AAAAAAATTA | CTAGGGTAGC | TAGTGATATT | GGTGTTAAGT | 3900 |
| | | 151000000 | | 1 mmccmc1 1 c | 100001110 | CAACTAAA | 2056 |

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5333 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 170:

| 10 | | | | | | | |
|----|------------|------------|---------------|------------|------------|------------|-------|
| 70 | ATTAAAACAA | CTTAATATAC | CTATTTATGG | TGGTCCTTTA | GCATTAGGTT | TAATCCGTAA | 60 |
| | TAAACTTGAA | GAACATCATT | TATTACGTAC | TGCTAAACTA | AATGAAATCA | ATGAGGACAG | 120 |
| 15 | TGTGATTAAA | TCTAAGCACT | TTACGATTTC | TTTCTACTTA | ACTACACATA | GTATTCCTGA | 180 |
| | AACTTATGGC | GTCATCGTAG | ATACACCTGA | AGGAAAAGTA | GTTCATACCG | GTGACTTTAA | 240 |
| | ATTTGATTTT | ACACCTGTAG | GCAAACCAGC | AAACATTGCT | AAAATGGCTC | AATTAGGCGA | 300 |
| 20 | AGAAGGCGTT | CTATGTTTAC | TTTCAGACTC | AACAAATTCA | CTTGTGCCTG | ATTTTACTTT | 360 |
| | AAGCGAACGT | GAAGTTGGTC | AAAACGTAGA | TAAGATCTTC | CGTAATTGTA | AAGGTCGTAT | 420 |
| | TATATTTGCT | ACCTTCGCTT | СТААТАТТТА | CCGAGTTCAA | CAAGCAGTTG | AAGCTGCTAT | 480 |
| 25 | CAAAAATAAC | CGTAAAATTG | TTACGTTCGG | TCGTTCGATG | GAAAACAATA | TTAAAATAGG | 540 |
| | TATGGAACTT | GGTTATATTA | AAGCACCACC | TGAAACATTT | ATTGAACCTA | ATAAAATTAA | 600 |
| | TACCGTACCG | AAGCATGAGT | TATTGATACT | ATGTACTGGT | TCACAAGGTG | AACCAATGGC | 660 |
| 30 | AGCATTATCT | AGAATTGCTA | ATGGTACTCA | TAAGCAAATT | AAAATTATAC | CTGAAGATAC | 720 |
| | CGTTGTATTT | AGTTCATCAC | CTATCCCAGG | TAATACAAAA | AGTATTAACA | GAACTATTAA | 780 |
| | TTCCTTGTAT | AAAGCTGGTG | CAGATGTTAT | CCATAGCAAG | ATTTCTAACA | TCCATACTTC | 840 |
| 35 | AGGGCATGGT | TCTCAAGGTG | ATCAACAATT | AATGCTTCGA | TTAATCAAGC | CGAAATATTT | 900 |
| | CTTAECTATT | CATGGTGAAT | ACCGTATGTT | AAAAGCACAT | GGTGAGACTG | GTGTTGAATG | 960 |
| 40 | CGGCGTTGAA | GAAGATAATG | TCTTCATCTT | TGATATTGGA | GATGTCTTAG | CTTTAACACA | 1020 |
| | CGATTCAGCA | CGTAAAGCTG | GTCGCATTCC | ATCTGGTAAT | GTACTTGTTG | ATGGTAGTGG | 1080 |
| | TATCGGTGAT | ATCGGTAATG | TTGTAATAAG | AGACCGTAAG | CTATTATCTG | AAGAAGGTTT | 1140 |
| 45 | AGTTATCGTT | GTTGTTAGTA | TTGaTTTTAA | TACAAATAAA | TTACTTTCTG | GTCCAGACAT | 1200 |
| | TATTTCTCGA | GGATTTGTAT | ATATGAGGGA | ATCAGGTCAA | TTAATTTATG | ATGCACAACG | 1260 |
| | CAAAATCAAA | ACTGATGTTA | TTAGTAAGTT | AAATCAAAAT | AAAGATATTC | AATGGCATCA | 1320 |
| 50 | GATTAAATCT | TCTATCATTG | AAACATTACA | ACCTTATTTA | TTTGAAAAAA | CAGCTAGAAA | 1380 |
| | ACCAATGATT | TTACCACTCA | ጥጥ አማር እ አርረጥ | 888CC88C88 | ****** | 202202220 | 2.446 |

. 55

50

| | GCTTTTTCTT | TATATATGAT | GAGCTTGAGA | CATAAATCAA | TGTTCAATGC | TCTACAAAGT | 1560 |
|----|------------|------------|------------|------------|------------|------------|------|
| | TATATTGGCA | GTAGTTGACT | GAACGAAAAT | GCGCTTGTAA | CAAGCTTTTT | TCAATTCTAG | 1620 |
| 5 | TCAGGGGCCC | CAACATAGAG | AATTTCGAAA | AGAAATTCTA | CAGGCAATGC | GAGTTGGGGT | 1680 |
| | GTGGGCCCCA | ACAAAGAGAA | ATTGGATTCC | CAATTTCTAC | AGACAATGTA | AGTTGGGGTG | 1740 |
| | GGACGACGAA | ATAAATTTTG | AGAAAATATC | ATTTCTGTCC | CACTCCCGAT | TATCTCGTCG | 1800 |
| 10 | CAATATTTTT | TTCAAAGCGA | TTTAAATCAT | TATCATGTCC | AATCATGATT | AAAATATCAC | 1860 |
| | CTATTTCTAA | ATTAATATTT | GGATTTGGTG | AAATGATGAA | CTCTTTGCCT | CGTTTAATTG | 1920 |
| 15 | CAATAATGTT | AATTCCATAT | TGTGCTCTTA | TATCTAAATC | AATGATAGAC | TGCCCCGCCA | 1980 |
| 15 | TCTTTTCAGT | TGCTTTCAAT | TCTACAATAG | AATGCTCGTC | TGCCAACTCA | AGATAATCAA | 2040 |
| | GTACACTTGC | ACTCGCAACA | TTATGCGCAA | TACGTCTACC | CATATCACGC | TCAGGGTGCA | 2100 |
| 20 | CAACCGTATC | TGCTCCAATT | TTATTTAAAA | TCTTTGCATG | ATAATCATTT | TGTGCTTTAG | 2160 |
| | CAGTTACTTT | TTTTACACCT | AACTCTTTTA | AAATTAAAGT | CGTCAACGTA | CTTGATTGAA | 2220 |
| | TATTTTCACC | AATTGCCACA | ATGACATGAT | CAAAGTTACG | GATACCTAAA | CTTTTCATAA | 2280 |
| 25 | CTGCTTCATC | TGTAGTGTCT | GCAACAACCG | CATGAGTAGC | GATATCACTA | TATTCATTCA | 2340 |
| | CTCTATTTTC | ATCATGGTCG | ATGGCCATTA | CATCCATGTC | TAATGCATTC | AACTCACGAA | 2400 |
| | CGATACTACC | TCCAAAACGA | CCTAGACCGA | TGACTACATA | TTCTTTACCC | ATACTCGCCC | 2460 |
| 30 | TCCATTAAAT | GATTTTCATC | AATTCATTGA | AAATATAAAT | TTAAAATTAT | TATAAATGAG | 2520 |
| | TACCCCAACT | AAATTATCTA | AATGCAGTAA | TGCAAGTAAA | TGAAAGTTGG | GGTATCGTCT | 2580 |
| | CAACTTATGA | TTTCTTTCCT | TCAACATATT | CTTTGTCGAA | AACAAATAAT | CTTAATAATA | 2640 |
| 35 | ATATTAACGA | TGGAAGTAAT | AAAAGTAAAC | СТААААТААА | GACAATCACT | AATGTCCAGC | 2700 |
| | CCAFTTCTGG | ATTAACATAT | GCATCTGTAA | TTTTTACAAA | CGGATATAAA | AGGTATGGCA | 2760 |
| 40 | ATTTACTAAT | TCCATAGCCA | AAGAACGCGA | ACATCATTTG | TAAAATAACA | AATACAAAAG | 2820 |
| | CCAAACCATG | TTTTTTCTTA | AAGAATGTTA | ACAATGAAGC | TAATGCAAAG | AATAAGAAAC | 2880 |
| | TTATACCAAA | CATCCACCAA | TAGTCAAAAA | CAGCTGAATA | AAAATGTTCA | GAATTTTGAA | 2940 |
| 45 | TGCGTAATGA | TAGAAATACG | AATAAACAAA | TGATAATCAT | CGGCGGCCCT | AAAAATATGT | 3000 |
| | GCCATTGTCT | TGTTAAATTA | TATGCTGGTT | CGTCATTTGC | TTTTTTAGCA | TAATATGTCA | 3060 |
| | AAAATCCTGA | TGAAATATAT | AAAACTGAAA | TAATTGCCAA | GAATACTACA | GACCAAGCAA | 3120 |
| 50 | ATGGGCTTAA | TAATAACTGC | ACCCAATCTA | GATCGATAAC | ATTGTTTCGA | ACATTAATAT | 3180 |
| | AGCCACETTC | TGTAATAGTT | AAAGCAGTAG | ATAATGAAGC | TGGAATTAAT | AATCCACTTA | 3240 |

| | AACTGTTTCT | CAACGATATC | ATAATCAGTG | CTATTGAACC | TGGTATTAAC | AATACCGTGC | 336 |
|----|------------|------------|--------------|---------------------|--|---------------------|------|
| | CTAAATATTT | GATTGACTCT | GGAAAGAAAC | CTACGAATCC | TACGAAGAAG | AAAACAAAGA | 342 |
| 5 | ATACATTCGT | AACTTCCCAA | ACTGGGTTTA | AATAACGTGA | AATTAAGTGA | TTAATTTTCT | 348 |
| | TTTCATCACC | AGTTAACTTT | GAATGCAATG | CGAAGAAACC | TGCCCCAAAA | TCTATAGAAG | 3540 |
| | CAATAATGAT | ATAGCAAAAT | AAAAACAACC | ATAACACTGT | TATACCTATA | AATGCATAAA | 3600 |
| 10 | TCATTTTTCT | ATTTCTCCTC | CTTGCTTCTT | GGCTAAACGA | TTTACATCTT | CATACGCCGG | 3660 |
| | TTTATTTTTA | AACATACGAA | TTAATACGTA | TGCACATGTA | TACATTAAAA | TGATGTACAA | 3720 |
| 15 | TATGCCAAAT | AAAATTGTAA | CGAaGGTTAT | TCCGCCTGCT | TGTGTTGCTG | CTTCTGCCAC | 3780 |
| | GCGCATATAA | CCACGAACAA | TCCAAGGCTG | TCTACCCATC | TCTGTTAAGA | ACCATCCAAA | 3840 |
| | TTCTATAGCT | AGCATTGAAG | CTGGGCCTGT | TAATAATATT | CCATAAAGCA | TCCATTTATG | 3900 |
| 20 | AGTAGAAAAC | TTTCTAAGCT | TTTTAAACAT | TAAAGTTAAG | ACATAAACAC | CTGAAATGAC | 3960 |
| | AAAACATAAA | ATTCCCATCG | TTACCATTAA | ATCAAAGAAA | TAATGGACGA | TCATAGGCGG | 4020 |
| | ATGTAAACTT | TTTGGAAAAT | CATTTAACCC | TTGTACTTTA | GTTTTGACAC | TATTATCTGC | 4080 |
| 25 | TAAGAAACTC | AATAGTCCAG | GTAATTCAAT | CGCACCTTTA | ACTTGCTGAG | TCTTTTCATC | 4140 |
| | TAACACACCA | AATAATAATA | ATTTGGCATG | GGAAGATGTA | TCGAAATGCC | ATTCATAAGC | 4200 |
| | TGCTAATTTT | TCAGGTTGGA | ATTTATGCAA | AAATTTTGCA | GATAAATCCC | CTGCCAACAT | 4260 |
| 30 | AGAAAGTAAT | GTTGAAAAGA | ATCCAACTAT | CATAGACATT | TTCAAAGCTT | TCTTATGGTA | 4320 |
| | GACAGTATCT | TTAGGTTGAC | GATTACGCAA | TAATTTAAAA | GCTGCTATTG | ATGCAATAAC | 4380 |
| | AAATGCCATC | GTCATACCGG | CTGTAGTAAT | TACGTGAAAT | GATCGAACTA | TAAACGATGG | 4440 |
| 35 | GTTAAACATC | GCTTCTATAG | GTTGAACATT | GACCATCTTT | CCATTCTTCA | ACTCAAAACC | 4500 |
| | TGCAGGCGTA | TTCATAAATG | AATTCACTGA | AGTAATGAAG | AATGCTGAGA | AAGAGCCACC | 4560 |
| 10 | AATAATTACT | GGTATACTAA | TTAAGAAATG | TGTCCATTTA | TTTTTAAAAC | GATCCCAAGT | 4620 |
| | TATAAATAT | ATACTTAAGA | AAATAGCTTC | AAAGAAGAAC | GCAAATGTTT | CCATAAATAA | 4680 |
| | TGGAAGTGCA | ATAACGTGTC | CACCCATTTC | CATAAATGTA | GGCCAAATCA | ATGATAATTG | 4740 |
| 15 | AAGTCCTATA | ATTGTACCTG | TAACAACTCC | CACTGCTACA | GTAATTGTAT | AAGCTTTAGC | 4800 |
| | CCATCTTTTG | GCCATAGCTA | TATATTGAAG | ATCATTTTTG | CGAATACCTA | AAAATTCTGC | 4860 |
| | AATTGCGAAC | ATTAAAGGCA | TACCAACACC | AATCGTTGCA | AAAATGATAT | GAACTGCTAA | 4920 |
| 60 | AGTCATAGCT | GTCAAAAACC | GACTGATTTC | AACTGTATCC | ATTTAAAAAC | ATCACCTTTT | 4980 |
| • | TCTTTTTTTG | ATGACAACAC | אידיים ממדתם | באיים א מידי מידיים | ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ | 773 777777777 5 5 5 | |

GAATTTCAAT GTATAATTGT GTATATTACA TTAGAATAAA GCACGAAGGA GCATGATACA 5160 TGTCAGAAAT AATCGTTTAT ACGCAGAATG ATTGTCCACC TTGTACATTT GTAAAAAATT 5220 ATCTAAATGA GCATCACATT GATTTTGAAG AGAGAAATAT CAACAATCAA CAATATCGAA 5280 ACGAAATGAT AGATTTTGAT GCTTTTTCAA CTCCGTTTAT TTTGTTGAAT GGC 5333 (2) INFORMATION FOR SEQ ID NO: 171: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 11126 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 171: ATACGTGACC CTTTATCCGA AAATTTCTTT TCATATTCTG TTAAAATATT ACTGCCATCG 60 TCTTCTTGAT GTAAATTTAG ATTTATTTTT GTAAAATACA TTCCAAATTG AGACATACTT 120 TCTAAACTGT AGGCAAATAG TCCTCTGTTA TCAGTTTTAA AATGTAAATC TCCTTCATCA 180 TTTAAGATTT GTTGATACAA CGCTAAAAAC GTATGATACG TTAAACGTCG TTTTGCATGA 240 CGATTTTTTG GCCATGGATC TGAAAAGTTC AAATAAATAC GCGAAACTTC GCCGTCTTTA 300 AAATATTCAT TTAATTCAAT GGCGTCATTA CAAATAATCT TTAAATTTGT TAAACCCATC 360 TCTTTAACTT TATCCAATAC TTTATAAACG ATACTTTTCT CACGTTCCAT TGAAATATAG 420 TTAATATGAG GATTTTGAGC AGCTAATGTT GTAATAAACT GCCCCATACC CGAACCAATT 480 TCAATGTGTA TCGGTTGCGT TTTATCAAAC CATTCAGTCA TTTTCCCTGC ATGTTGACCG 540 TCCATGTCAA CCAATTCAGG ATGATCTTT AAATAATCTT CAGCCCATGG TTTGTATCGA 600 ACTCTCATAT TTTATTCTCC TCTTAAATAA ACATGTTACT ATTCATAACT TCATTTAGGA 660 ATTTAAGCCA AGTGTTCATA TCCTTATATC TTTTTTGCTC TTCATACCAT TGAACAAGAC 720 CTATAGATTG AATTACCGTA TACCATTTCA TACGTTTATT TAAATTCAAG CTCTCTTGAA 780 CACCATATGT TTCAAGCCAT TCAGACCATT GTTGTTGTGG AACATAGTTG TAAAGCAGCA 840 TTCCGATATC AATTGCCGGG TCTGCAATCA TTGCACCTTC CCAATCAACT AAAAATAGTT 900

CATCTCGATC GGATAATAAC CAATTATTAT GATTCACATC ACCATGTACA ACAGTGAAAA

AACGCGAATC TAAACTCGGT ATATGCTCTT CTAAATAGGT TAATGATTTT CTCACAATAT

GATGTGTTAA AACTTCTCTT GATAAAGAGG CATTAATTTT ATTAAGCATA ATCTCAGGAG

TAATAGGTTC CATTTCCATA CGCTTTAACA TACTTAATAA AGGTCTAGAA TTGTGTATCT

55

50

5

10

15

20

25

30

35

40

45

960

1020

1080

| | TTTTCCAATO | TTGTGCTGTA | ACAACCTCGC | CIGITICIA | GCGTTTCGT | CATACTAATT | 126 |
|-----------|------------|------------|------------|------------|------------|------------|------|
| | TGGGCACAAT | ACCTTCTGCT | GATAATGCCG | CAATAAATGO | ATTTGAATT | CGTTTTAAAA | 132 |
| <i>5</i> | ACAACTTTTG | TCCATCTTGT | TCAGCCATAT | ATGCTTCACC | AGATGCACC | CCTGCTGAAT | 138 |
| | CAAGTGTCCA | CCCTAATTGA | TAAAACTGCT | CCAACTCGTC | CACCTCACTI | TCAATTAGAA | 144(|
| 10 | AATGGCTCTA | GAAATAGGTT | TTTCAAGAGC | CATATATTCT | AATTTATAAC | ACCATACTGG | 1500 |
| 70 | TACAAATATT | ATGTCCAGAT | AATTATTGTA | AATCCTCAAC | CAATGCCTAC | ATTACACGAC | 1560 |
| | TAAATTTAAA | TCGTAATGTC | TGTCATTGAC | ACCATACATT | CTATAGTCAC | TTACTTGACA | 1620 |
| 15 | TATAATGTTA | CCGTGTCTAA | AACTACATGT | TTTTGAATCT | CTGTAGGCGA | TAAACTCTAG | 1680 |
| | TTTTCAAAAT | AATTGCTATC | CCATTTTCAT | GGTTAGCATA | AATTTATGAA | CTGTAACATT | 1740 |
| | TACGTACTTA | GTAAAATATG | ATGCACATCA | TATTTGTTAC | TCATAGAAAA | TTTTATAALT | 1800 |
| 20 | TTTATCATTA | TATTTCAACT | GAAAATGAGA | AACAAAATGG | CACTTTTTAC | TAATATGTGT | 1860 |
| | TTTCTAAACA | ACACTTTTAA | GCTTCGTTTT | AAATTATAAC | ATAATTCACT | TACGAAAGTT | 1920 |
| | GATAAATTTA | AGTAATTTAA | тстааааата | TGATGAAAGA | ATTTTAAATA | CTGTGTGACT | 1980 |
| 25 | CTATATACTT | TTCAAATCCT | TCTTGTAGTT | GACGTGTAAT | TGGGCCAACT | TTACCATCAT | 2040 |
| | TAACTGGTTC | ACCATCTAAT | TTAATAACAG | GTGTAACCTC | AGCTGAAGTA | CTTGAAACAA | 2100 |
| | TAACTTCATC | TGCGTTTTTC | AAGAAATCTA | CAGTAAACGT | TTCTTCTTTA | AATGGGATGT | 2160 |
| 30 | TATAGTCTTC | GGCAATTTTT | TTAATTACAA | TTCGTGTAAT | ACCATTAAGA | ATATAGTTGT | 2220 |
| | TAATCGGATG | TGTATAAATC | ACACCGTCTT | TAATTGCATA | AGCATTACTT | GAAGATCCTT | 2280 |
| 25 | CAGTTACAGT | TtCACCTCGA | TGTTGAATTG | CTTCAACTGC | ATTATATTTC | ACAGCATATT | 2340 |
| 35 | CTTTTGCTAA | TACATTCTCC | TAATAAGTTC | AAGCTTTTAA | TGTCGCAACG | TAACCATCGG | 2400 |
| | ATATETTCAA | CGGTAACACC | ATTCACACCA | TTTTCTAAAT | GATCATAAGG | ACGATCATAA | 2460 |
| 40 | CTCTTTGTAT | AAGCAACAAT | TGCTGGTTCT | ACTTCAGGTG | TCGGGAAGCT | ATGATTCCTT | 2520 |
| | TCAGCTACAC | CACGCGTTGC | TYGAATATAA | ATTGCCCCAG | TTTCAATTTG | ATTCATATCA | 2580 |
| | ACTAATTTÄC | GAGATAGTTC | AATTAATTCT | TCTACAGAAT | AATTTAAATC | TAAACCAATC | 2640 |
| 15 | TCATTGGCAC | TACGTWAAAW | TCTTTCATAA | TGTTCTGTTA | CTGTAAATAA | CTTACCATTA | 2700 |
| | TATACTCGAA | TGTATTCATA | AATACCATCG | CCAAATACGT | ATCCTCTGTC | GTTGTATGAA | 2760 |
| | ACCTTTGCTT | CACTTGGACT | TACAAACTCA | CCATTTAAAA | AAATTTTTTC | CATATATTAT | 2820 |
| 50 | TCCTCCACGC | ATAATGAATA | AATTGCTTCT | AAGTAAATAC | TAGTTGCGTT | AAATAACTGT | 2880 |
| | TTTTTAGTGA | TATATTCATT | TTTCTGATGC | ATTAAATCTT | САСААТСАСТ | AAACATTGCG | 2940 |

| | TCAGTCATAT | CATTTGTTTG | ATTTCTATAT | GCAGTAACTA | ACTTTTGTAC | AAAAGGATCA | 3060 |
|----|------------|------------|------------|------------|------------|------------|------|
| | TTTTTATCAA | CATAATGTGG | TGGTTGGACT | TTACCTAATT | TCACTTCAAA | GCCATATTGT | 3120 |
| 5 | TGAATCTCAT | TTGCAAAACG | ATCCATAGCT | TTTTCAAATT | CAAATCCTTC | TGGGTAGCGT | 3180 |
| | AAGTTGATAC | CGAAAAGACC | TGCGTTTTCA | TTATCATATG | TAATAACACC | AATGTTAGTT | 3240 |
| _ | GTCACGTCAC | CCATGACATC | TGTATGGAAT | TTCATTCCCA | TCTTTTCACC | AAAATCTGAA | 3300 |
| 0 | TTAAATAAGT | AGCGATTACT | AAATGCTACA | AACGCTTGTG | CATTATTATC | AAGATTTAAT | 3360 |
| | GATGCTAAGA | ATTTTAGTAA | GTAAAGACCC | GCATTCACAC | CGATAGATGG | ATCCATACCA | 3420 |
| 5 | TGAACCGCTT | TACCTTCAAC | TGTTAAAACT | AGAATGCCAC | TATCAACAGT | ACTATCACCT | 3480 |
| | TGTAAATGAT | TTTGTTCTAA | AAAGTACTCA | AAGTCTTGAA | TAACATCTGT | CATATTTTCT | 3540 |
| | TTAACAAGCA | CTCTTGCTTC | TGCATGATCA | GGTACCATGT | TGTAACGTTC | ACCAGATTTA | 3600 |
| 20 | AAAGTTATTA | ATTCATAATC | AGGTTCATCT | TGATCTTCAG | TAAGTTTATT | TTGAACTAAA | 3660 |
| | TCAAATGTTG | TAATGCCTTT | TTCACCATGA | ATACATGGAA | ATTCTGCATC | TGGTGCAAAA | 3720 |
| | CCTAATGTTG | GCATTTCTTC | TGTTTTAAAA | TAGCGATCCG | TACATTTCCA | ATCAGATTCT | 3780 |
| 5 | TCATCCGTAC | CAATAATCAT | ATGAATACGT | TTCTTCCAAT | CCACATTCAT | ATCTTCTAAT | 3840 |
| | ATCTTAATTG | CATAATAAGC | AGCAATTGTT | GGACCTTTGT | CATCAAGTGT | ACCTCTAGCT | 3900 |
| | ATGATAGCAT | CTTCTGTTAC | AACCGGCTCG | AACGGATTAC | TATCCCATCC | ATCACCAGCA | 3960 |
| 0 | GGAACAACGT | CAACATGACA | TAAGATACCT | AATACGTCAT | TTCCTTTACC | TGCCTCAATT | 4020 |
| | CTTCCTGCAA | TATGATCCAC | ATCATGTGTT | GTAAATCCAT | CTCTATGTGC | AATTTCATAC | 4080 |
| | ATGTAGTCTA | ATGCCTTACG | AGGACCTGGA | CCAACTGGTG | CGTCTTCTGA | TGCTTTTGCA | 4140 |
| 5 | TCATCTCTCA | CACTTTCAAT | TGCTAATAAT | CCTTTTAAGT | CATTAATGAT | TTGATCTTCG | 4200 |
| | TATTGTTGAA | CTTTTTCTTT | CCACATTCGA | AATCGACTTC | CTTTTTTCTA | TAAGTTAAAT | 4260 |
| o | TCTATTTTAC | ATGAAAAGAT | АТАААААСТА | CAATAAGATG | TCAGAAAATA | ATAAAAAGGA | 4320 |
| | ACAAAACGAT | GCTATTGATA | TGACACAAAT | CATAAATAGC | TGCTTTGTTC | CTTTTTTAAT | 4380 |
| | TTATATATT | AAAATACACA | TATTCAAGAG | CTCGAGATAT | AAGTCAATGT | ACTAGGCACA | 4440 |
| 5 | CAATTTAATA | TTGACAGTAA | TTAACCGAAC | GAAAATGCGC | cccccccc | CAACATAGAG | 4500 |
| | AATTTCGAAA | AGAAATTCTA | CAGACAATGC | AAGTTGGCGG | GGCCCCAACA | TAGAAGCTGG | 4560 |
| | CCAATAGTTA | GCTTTCAATA | ATGTGCAAGT | TGGGGTAAGG | GCCCCAACAC | AGAAGCTGGC | 4620 |
| 0 | CAATAGTCAG | CTTTCAATAA | TGTGCAAGTT | GGGGTAAGGG | CCCCAACACA | GAGAATTTCG | 4680 |
| | ΑΑΑΑΩΑΑΑΤΤ | СТАСАСАСАА | TGCAAGTTGG | CCCCCCCA | ACACAGAAGC | TCCCCAATAG | 4746 |

| | TAAAGAAATA | CGTTTTCTTT | AGATATTAGT | ATTTCTTATG | AATGAGTTTC | ACGCATGTAT | 4860 |
|----|------------|------------|------------|------------|------------|------------|------|
| | TCTTCTTTCT | ATATGCATAT | TAGCTATGAC | TAACGATAAA | GAACCTGAAA | CACTAATAAA | 4920 |
| 5 | TGTCCTATAG | TTTACAATAT | TATATTGGCA | GTAGTTGACT | GAATGAAAAT | ACGCTTGTAA | 4980 |
| | CAAGCTTTTT | TCAATTCTAG | TCAACCTTGC | CGGGGTGGGA | CGACGAAATA | AATTTTGCTA | 5040 |
| | AAATATGATT | TCTGTCCCAC | TCCCTTATCA | тттстстсст | ACTCACATCT | TATTCTTTAT | 5100 |
| 10 | CAGATAATGC | ATTTTTATTC | TTTTTTAAAT | CTTCTTCAGT | GACGATACGT | AAATTATTAT | 5160 |
| | TTGGTGTGCG | CCACCTTCAT | CATCAAATTT | ACCTTTTTCA | ATACTTTCGT | CAGTCTTATT | 5220 |
| 15 | GTCATATTCG | GTAAATTTTG | ATTTTTCTTC | TTTGAAAAAT | GCTTTTGGAT | TATTTTTTAA | 5280 |
| | TCTATTAGCA | TATTCTTTCG | GATTTGTTTT | TACTTCTTTA | ATTGTTTCAT | TAGCAATTGT | 5340 |
| | TCCTAATTGC | GTCGCTTTAT | CCTTAGCATT | ATCTTTATAG | CTTTGAGGAT | CTTGTTTATA | 5400 |
| 20 | TTTATTATAT | TCcTGcTTTC | AGCTTGTCAC | GACTATCTTT | ACGTGTAACA | AGTACAGCTG | 5460 |
| | CTACAGCGCC | ACCTATACCT | AAAATCGCTT | TAAATAAATT | ACCTTTTGCC | ATATCAATCG | 5520 |
| | TCTCCCTTTT | ATTTATAATT | TAATTTGTCA | AAATCATTTT | CAGTTAATAA | ACGATATTCT | 5580 |
| 25 | CCTGAATCTA | AATTGCTGTC | CAATTCTAAA | TCAGCAATTT | TGATACGTCT | TAAATGTAAT | 5640 |
| | ACCTCATTTT | GAATGCTATG | AAACATTCGT | TTAACTTGAT | GATATTTTCC | TTCATAAATT | 5700 |
| | GTTACGTGTG | ACGTTTGATT | ATCAATATAA | GTTAATATTG | CAGGCTTAAC | CTTGCCATCA | 5760 |
| 30 | GTCAGTGTtA | CACCCTCTTT | AAAAGCTTGA | ATGTCGTCTT | CAGTGATAGG | ATTTGCTGAA | 5820 |
| | ATAACTTCAT | ATTTTTTAGA | AACATGTTTG | TTTGGACTCA | TTAATTCATG | ATTAAAATCA | 5880 |
| 35 | CCATCATTCG | TTATCAATAA | AAGCCCTTCT | GTATCTTTAT | CAAGACGACC | AACCGGAAAA | 5940 |
| | ATATTTAGAT | GTTGGTATTC | AGGTATTAAA | TCAATAACGG | TTTTTGAATG | ATGATCTTCA | 6000 |
| | GTTGCTGATA | TATAACCTTT | TGGCTTATTT | AACATAATAT | AGACATTTTC | AATGTATTCT | 6060 |
| 40 | ATTAATTCTC | CACGAACTGT | TATCTTATCG | TTTTCTGGTT | CTATATGTGT | TTTTGGTGAT | 6120 |
| | TTAATTACTT | GTTCGTTGAC | ATTTACAAGG | CCTTTTTTAA | GTAACTGTTT | GACCTCATTA | 6180 |
| | CGTGTACCGA | CGCCCATATT | TGCTAAAAAT | TTATCTATTC | TCATCGTAAA | AACCTAACTC | 6240 |
| 15 | TACGTCTTAA | TTTTTCAGGA | ATTTCACCTA | AGAATTCGTC | CGCAAGACGC | GTTTTAATTG | 6300 |
| | TGATTGTACC | GTAAATTAGA | ATACCTACTG | TAACACCTAA | AATAATAATG | ATTAAGTAAC | 6360 |
| | CAAGTTTAGT | AGGTTCTAAG | AATAGATTTG | CAAGGAAAAA | TACTAATTCT | ACACCTAGCA | 6420 |
| 50 | TCATAATAAA | TGAATACAAG | AATATTTTTG | CAAAATGAAT | CCAACTATAG | CTGAATTTAA | 6480 |
| | ACTTCGCATA | THEFT | ATATAGAAAT | TACATCCAAT | ጥርርልልልጥልልጥ | AATCCCATAC | 6540 |

| | ACTTGATAAC | TACAGAAGCT | AAAATAACAT | AAACTGTTAA | TTTCTGTTTA | TCTATACCTT | 6660 |
|------------|---------------------|------------|--------------|------------|------------|------------|------|
| | GTAACATTGA | TGCCGTTACA | CTTAATAGTG | AAATTAGTAT | TGCTACAGGC | GCATAATAGA | 6720 |
| 5 | ATAATAAGCG | ACTACCATCA | TGGTTAGGGT | CATGACCTAA | AACAATTGGA | TCGTAACCAT | 6780 |
| | AGAAAACTGT | GAATAATGGT | TGTGCCAAGG | CCATAATTCC | AATACTAGCT | GGAACAGTTA | 6840 |
| | TAAACATTAA | TACACCAATA | GATGTTCTAA | TTTGATGATG | CATTTCATGT | AAGCGACCTT | 6900 |
| 10 | CTGCAAATGT | TTTTGTAATA | TAAGGAATTA | AACTCACTGC | AAAACCAGCA | CTTAATGATG | 6960 |
| | TCGGAATCAT | TACAATTITA | TTAGTTGACA | TATTTAGCAT | ATTAAAGAAT | ATATCTTGTA | 7020 |
| 15 | ACTGTGAAGG | TATACCAACT | AAAGATAAAG | CACCGTTATG | TGTAAATTGA | TCTACTAAGT | 7080 |
| | TAAATAATGG | ATAATTCAAA | CTTACAATAA | CGAACGGTAT | ACTATAAGCA | ATAATTTCTT | 7140 |
| | TATACATCTT | GCCATATGAC | ACATCTATAT | CTGTGTAATC | AGATTCGACC | ATACGATCAA | 7200 |
| 20 | TATTATGCTT | ACGCTTTCTC | CAGTAATACC | AGAGTGTGaA | TATACCAATA | ATCGCACCAA | 7260 |
| | CTGCTGCTGC | AAAAGTAGCA | ATACCATTGG | CTAATAAAAT | AGAGCCATCA | AAGACATTTA | 7320 |
| | GTACTAAATA | ACTTCCGATT | AATATGAAAA | TCACGCGTGC | AATTTGCTCA | GTTACTTCTG | 7380 |
| 25 | ACACTGCTGT | TGGCCCCATA | GATTTATAAC | CTTGGAATAT | CCCTCTCCAT | GTCGCTAATA | 7440 |
| | CAGGAATAAA | GATAACAACC | ATACTAATGA | TTCTTATAAT | CCAAGTAATA | TCATCGACTG | 7500 |
| | ACCAACCGTT | TTTATCATGA | ATGTTTCTAG | CTAATGTTAA | TTCAGAAATA | TAAGGTGCTA | 7560 |
| 3 0 | AGAAATACAG | TACCAAGAAA | CCTAAAACAC | CGGTAATACT | CATTACAATA | AAACTCGATT | 7620 |
| | TATAAAATTT | CTGACTTACT | TTATATGCCC | CAATAGCATT | ATATTTCGCA | ACATATTTCG | 7680 |
| | AAGCTGCTAA | TGGTACACCT | GCTGTCGCAA | CTGCAATTGC | AATATTATAT | GGTGCATAAG | 7740 |
| 35 | CGTATGTGAA | CGGCGCCATA | TTTTCTTGTC | CACCAATTAA | ATAGTTGAAT | GGAATGATAA | 7800 |
| | AAAGTACGCC | CAATACCTTG | GTAATTAATA | TACTAATGGT | AATTAAAAAG | GTTCCACGCA | 7860 |
| 40 | CCATTTCTTT | ACTITCACTC | ATTACGAATC | TCCCTATCTC | ATGTTTATTA | AAGTTTTGTA | 7920 |
| •• | AACTAAAAGC | TGTTTCTCTG | TAAAATCATT | TTTCATTATT | ATGAATATAT | CACAAAACTT | 7980 |
| | TATTTCATTG | TCGTATATTC | AATGAATTAT | CATAACAAAA | TTATCAACAC | ATTGTCATTG | 8040 |
| 45 | AATACTAGAT | TTTGATTAGA | ATATTACGAA | ATTTCATATA | AACATTATAC | TACTATTTGA | 8100 |
| | GATGAACATC | GCATAACAGT | AGAAAAATCA | TTCTTATCAT | ACACATACAT | CTTCATTTTT | 8160 |
| | TATGAAGTTC | ACATTATAAA | TATATTCAAC | ATAATTGTCA | TCTCATAACA | CAAGAGATAT | 8220 |
| 50 | AGCAAAGTTT | AAAAAAGTAC | TATAAAATAG | CAATTGAATG | TCCAGTAACA | AATTTGGAGG | 828 |
| | 3 3 C C C T 3 T 3 T | CT3TC333C3 | ************ | Checesare | TACCCCCTTA | ATCCCCCCAC | 8341 |

| | GTAAACTCA | A AATATCTGGT | GGCGGTAGA: | r gtaacgtaac | TAATCGATT | A CCATATGCTG | - 8460 |
|----|------------|--------------|------------|--------------|------------|--------------|-----------|
| | AAATTATTA | A GAACATTCCT | GGaAATGGG! | ATTTTTAA A | TAGTCCCTT | r TCAATTTTTG | 8520 |
| 5 | ATAATGAAT | CATCATAGAT | TTTTTTGAG | CTAGGGGTGT | TAAATTAAAI | A GAAGAAGATC | 8580 |
| | ACGGGCGTA: | r gtttccagti | TCCAACAAA | G CACAAGACGT | GGTTGATAC | TTAGTGACAA | 8640 |
| | CTATCGAAC | G CCAACATGTA | ACGATTAAAG | AAGAAGAAGC | TGTTAGTAG | ATCGAAGTTA | 8700 |
| 10 | ATACAGACC | AACTTTCACT | GTACATACTO | AAAATAATAG | TTATGAAAGC | CATTCGCTAG | 8760 |
| | TGATTGCTAC | AGGTGGTACA | AGTGTCCCTC | AAACTGGTTC | AACTGGTGAT | GGTTATAAGT | 8820 |
| 15 | TCGCACAAGA | TTTAGGTCAT | ACCATTACTG | AGTTATTCCC | GACCGAAGTT | CCAATTACAT | 8880 |
| | CAGCTGAACC | TTTCATCAAA | TCCAATCGTC | TAAAAGGTTT | AAGTTTAAAA | GATGTTGAAT | 8940 |
| | TGTCAGTACT | TAAGAAAAAT | GGTAAAAAAC | GCATCAGTCA | TCAAATGGAT | ATGTTATTTA | 9000 |
| 20 | CTCATTTTGG | TATCAGTGGT | CCAGCTGCAT | TAAGATGTAG | TCAGTTTGTT | TATAAAGAAC | 9060 |
| | AAAAAAATCA | AAAGACACAG | CACATTTCTA | TGGCAATCGA | TGCATTTCCT | GAATTAAACC | 9120 |
| | ATGAACAATT | AAAACAACAC | ATCACATCAT | TATTATCGGA | CACACCAGAT | AAAATCATTA | 9180 |
| 25 | AAAACAGTTT | GCATGGTCTA | ATTGAAGAGC | GCTACTTACT | GTTCATGCTG | GAACAAGCAG | 9240 |
| | GAATCGATGA | AAATACCACA | TCACATCACT | TATCAAATCA | ACAATTGAAC | GACTTAGTAA | 9300 |
| | ATATGTTTAA | AGGGTTTGTA | TTTAAGGTGA | ACGGGACATT | ACCTATAGAT | AAGGCATTIG | 9360 |
| 30 | TCACAGGTGG | TGGTGTGTCA | CTTAAAGAAA | TTCAACCTAA | AACAATGATG | TCTAAATTAG | 9420 |
| | TTCCGGGATT | ATTTTTATGT | GGTGAAGTAT | TAGATATACA | TGGTTATACT | GGTGGTTATA | 9480 |
| 25 | ATATTACAAG | TGCACTCGTA | ACAGGACATG | TCGCTGGATT | ATATGCCGGA | CATTACTCAC | 9540 |
| 35 | ATGCATCAAT | GGAATAATAG | TATAAAATTT | GGTTCGATTC | TCTTTAGTAG | ATCAACTTTT | 9600 |
| | TCATTCAAAT | AAAAATGACC | TTAATATAAC | TGAGTCACTA | AAAAGTGTCG | TTATATTAAG | 9660 |
| 40 | GTCATTTCGT | TAATTATGAT | TCTTTTTCGT | TTTTAGTACG | TCTTCTAGCT | AACAAAGCCG | 9720 |
| | CACCTGTAAT | CAGTGCAAAT | TCTTTCAATG | GTAAATCCAT | TCCTTCAGAA | CCTGTATTTG | 9780 |
| | GAAGTTCTTT | TTCAACTTTG | CGCGATTCAT | GTGTCTCTTC | TTTTTTAATA | GGCGTACAAA | 9840 |
| 45 | CTTTTGGAGC | TGGCTGAATT | TCTTTTGGTG | ATACTTTCGT | CGCTTCAGCT | GGTAATTTAA | 9900 |
| | TTGCTAAAAT | TTCATCAACA | ATGAATTGCG | TGTGTTGTTT | GATGTCATTT | AATGTCGCAT | 9960 |
| | CTTCATCAAT | CATTCTATTG | CCATCTGCAA | CATATTGATC | AATTAATACT | TTTACTTTAG | 10020 |
| 50 | CTAATTGTTC | TGGTGTTGCG | ATCGCTTTGA | ATTTCGCATA | TGTTTGTTGA | GCAATGTTAT | 10080 |
| | CAATTCGCAG | TAAGCTATTT | TCTTTTTCAG | TAATTACTICC | ттстататсс | СТТААТССАЗ | 10140 |

| CATCCATTTG | TAATTTTAAA | GCAGTTATAG | CTTTTAATGC | ATCAGCCTTA | TTACGATTAC | 10260 |
|------------|------------|------------|------------|------------|-------------|-------|
| TTACTTTTCG | ATAATTTTGC | ACTAAAGCAG | TGACGCGTGC | AAGATCATCA | TTAATCGTTT | 10320 |
| TTTCAGCATC | TGGCTTTTTA | ATAGGATGTA | CATCTAAATC | ATGTATTGTT | TGTAGATTTA | 10380 |
| ATGATGCTGT | TTTATCAACT | TGTGCATTGC | TACGATCTTG | ATCAATTIGT | CCAATAGCAG | 10440 |
| TGTCATAAAT | ATTTTGTAAC | TGTGCTAATA | TACTATTTCT | TTCTTCTACC | GTTGCTTGAA | 10500 |
| TATTCGCTTC | AATTGCTTGT | TTTTTATCGT | TGAATAATGT | TGTCAATTGT | TCTCGAGCAG | 10560 |
| ACGCCTTTCT | GTTAATAACA | GGTTCGATTT | CACGAATTTC | GTTTTTCTCA | TCATGCAATA | 10620 |
| AATATGCCAC | ATCTGCATTA | GTCACTGCAC | TAGCAATTTG | TTGTTTAGCT | TTAATTAACT | 10680 |
| CTTTTTCAAC | TTGTGCTATT | GCAATATTTT | GTTCTTCATC | TGTCGCTTCG | TTATTTGCTT | 10740 |
| TAATTAAATT | AATTTTATTT | GTAGCGATAT | TTTGAATTTG | TTGTAATGCT | GTTGCTTTAA | 10800 |
| CTGTTGTCGC | TGGTTTAATT | TTTGAAATAA | TATTTTGAGC | ATTTATACTA | TCTTGATTAA | 10860 |
| CTTGGGCAGT | CTTATCTGCA | TGATTGATCT | GATCAATAGC | CTGATTAAGT | GCTTGTTCTA | 10920 |
| CTAAATGTTT | AGCAGCTAGT | CTTTCTTCTT | CAGTTGATAA | ATCGCTTTGA | TCGATTAGTG. | 10980 |
| CATTTTGAGC | TTCGGCTTTT | ACACCAACAG | ATTGACGCGC | TGCTGGTTTA | ACTTGAACTT | 11040 |
| TAGGTAAAAT | CACTTTGATG | TTGTCGTTGC | CATCAGTCnC | AGTnCGATCC | ACTTCTGCAT | 11100 |
| TCGTTTTGTT | TTGTGCAATG | TCATTT | | | | 11126 |

(2) INFORMATION FOR SEQ ID NO: 172:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3660 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 172:

TTGCCCCGCA CGGCGGTGTG nTTCCTAGAA ATAATGAATA TAAAGAGAAA TATATAACAA 60
CGATTTTGAA TTATGAACCT GGTGATATCG TTACAATCAA ACGTGTGAGA GATAAGACCG 120
ATTTGCTAAT ATATTTGTCT AGTAAAGATA TTTCTATTGG TAATGAAGTG GAAATTGTAT 180
CGAAAGATGA AATGAATAAA GTAATTATCA TTAAACGTAA TGATAATGTA ATTATTGTCA 240
GTTACGAAAA TGCAATGAAC ATGTTTGCTG AAAAATAAAA TAAAGAAGCC ATAAAGATAT 300
CCATGATTGA ACTGATAAAG ACATATGGAT AATTGCTTTA GGCTTCTTTT TTATTAGTTA 360
ATTTATCAAG TGAGTATATT TGAGTAAAAT ATTCACTGCA TAAAGATTGA AGATAATCCA 420

55

5

10

15

20

25

30

35

40

45

| • | CTGTGGACTC | GGACGCTGGA | AAGTCAATTT | AGCAATCGTC | CAACTAGATT | GTAGAACTTC | 54 |
|-----|---------------------|--------------------|------------|-------------|------------|----------------------|------|
| | GCCTAATAAT | ACACCTAAAA | TATATTGATA | ACTCATTGTG | ACAAGTAGTT | GAATTTCTAC | 60 |
| 5 | TATATTTTCA | TCTTTTAATA | TAAAATACAA | CATGATAGAA | ATTAAAGTTA | TAACAACAAT | 66 |
| | GGGTGAGCCT | TTTCtAGATG | TTAAAATTAA | AAAATÁAATA | AATATCAATA | AATAGGTAAA | 72 |
| -10 | TATAAAGAAA | CTAGGTATCT | GATAATGGCT | CGACGCTAAA | CCTATCAATA | ACATAATAGG | 786 |
| 10 | TGGCATAAAA | TAACCACCAA | TCGTTGTAAG | CCATTGGCCT | GCTAGATGTC | TAGATTGTGT | 84 |
| | AATTGCGAAT | CCTTGTTGTA | ATGTCTGTTG | TCGCTCTCGT | GGACTTGTTA | CAATGACTAA | 900 |
| 15 | ATCTTTTGCA | CGGCCACCAG | CGAGTTTATT | AAACAGTACA | TGACCAAATT | CATGTGTTAA | 960 |
| | AACAGGGATA | TAGTTTAAAA | TGACATCTAA | ATAGTTCAAA | ACAGGCTTAT | GTCTATATTG | 1020 |
| | ATGAATAGCA | ATATAACAAG | CTGCAACAAT | AACGATAATG | TATATATTAA | GTTGAATTGT | 1080 |
| 20 | CGTATTAAAA | AAGTTTGATA | AATAATTCAT | TGTTAACCTC | ATATAAGATA | TTAATTTAAA | 1140 |
| | GTTTGCTTAT | CACTTATTAT | AAATGATATT | GGCATCAATA | GCGTTAGACT | TTAGACTTAC | 1200 |
| | CTTAGTTAAA | CTAATTTTAA | TTTTTGAAAA | GGTGAATATG | TGTTAAAATA | AAGCAAAATC | 1260 |
| 25 | ATTTCGATAT | AAATAGGATG | AATATAAATA | CTGTTAATAT | TGATTACACT | AACATAATAA | 1320 |
| | TGAAATAAGA | TAGGAGATTC | CTGTTATGAC | TGTTGAAGAA | AGATCCAATA | CAGCCAAAGT | 1380 |
| | TGACATTTTA | GGGGTCGATT | TTGATAATAC | AACAATGTTG | CAAATGGTTG | AAAATATTAA | 1440 |
| 30 | AACCTTTTTT | GCAAATCAAT | CAACGAATAA | TCTTTTTATA | GTAACAGCCA | ACCCTGAAAT | 1500 |
| | AGTGAATTAC | GCGACGACAC | ATCAAGCGTA | TTTAGAGTTA | ATAAATCAAG | CGAGCTATAT | 1560 |
| | TGTTGCTGAT | GGGACAGGAG | TAGTCAAAGC | TTCGCATCGT | TTAAAGCAAC | CTCTAGCGCA | 1620 |
| 35 | TCGTATACCT | GGTATTGAGT | TGATGGATGA | ATGTTTGAAA | ATTGCTCATG | TAAATCATCA | 1680 |
| | AAAAGTATTT | TTGCTAGGGG | CAACTAATGA | agttgtagaa | GCGGCACAAT | ATGCATTGCA | 1740 |
| 10 | ACAAAGATAT | CCAAACATAT | CGTTTGCACA | TCATCACGGT | TATATTGATT | TAGAAGATGA | 1800 |
| | GACAGTAGTG | AAcGnAnTTA | AACTGTTTAA | ACCTGATTAC | ATATTTGTAG | GTATGGGATT | 1860 |
| | CCCTAAACAA | GAAGAATGGA | TTATGACACA | TGAAAACCAA | TTTGAATCTA | CAGTGATGAT | 1920 |
| 15 | GGGCGTAGGT | GGTTCTCTTG | AAGTATTTGC | TGGGGCTAAA | AAGAGAGCGC | CTTATATCTT | 1980 |
| | TAGAAAATTA | AACATTGAAT | GGATATATAG | AGCATTAATA | GATTGGAAAC | GTATTGGTAG | 2040 |
| | ATTAAAGAGT | ATTCCAATAT | TTATGTATAA | AATAGCCAAA | GCaaaaagaa | AAATAAAAA | 2100 |
| ю | GGCGAAATAA | TCATGATGAC | AAAAATAAAA | CCGAGGAAAT | CCTTAAATGG | AGATTCTCGG | 2160 |
| | للماسلسليل (والأناء | מידה מידיידי מידיד | ACCARCCCC | A CTCATCCAC | | A B COTTONION TO THE | 222 |

| | CATCAAGTTC | ACCGTAATCT | TTTAACTTTC | CGCCTTCAAT | CCAAGCAATC | TTAGTACAAA | 2340 |
|---|-------------|--------------|-------------|------------|------------|------------|------|
| | ATTGTCTCAC | TTGTCCTAAG | TTATGACTAA | CGAAAAAGAT | GGTTTTGTTT | TGCTCTTTAA | 2400 |
| 5 | ACTCGTAAAT | TTTATCTAAA | CATTTTTGTG | CAAAAGTTTG | GTCACCTACA | GATAAAGCTT | 2460 |
| | CGTCAATGAC | TAAGATATCT | GGATTAACTG | TGATATTAAT | TGAAAAACCA | AGTTTTGCAC | 2520 |
| | GCATACCACT | TGAATACTTT | TTAACTGGTT | GATAAATAAA | CTCACCAAGT | TCACTAAATT | 2580 |
| 0 | CAATAATCTT | AGGTGTCATC | GCTTTAATTT | CTTTTCGCTT | AAAGCCCATA | CATAACATTT | 2640 |
| | TAAATTCGAT | ATTTTCAATC | CCTGTAAGTT | GTCCACTCAA | GCCAGCACTA | ATTGCGATAA | 2700 |
| 5 | CGCTGACTTC | ACCATTACGA | TCCACTTTGC | CAACAGTAGG | CGACAAAGAA | CCGCCAATGA | 2760 |
| - | TATTGCTCAA | CGTTGATTTG | CCGGAACCAT | TGATGCCAAC | AAGCCCTATG | ACGTCACCTT | 2820 |
| | CATATGCTTT | TAAACTAATG | TCATCTAAAG | CGAAAAATGT | TTTGTTTTTA | TGTTTGGGAA | 2880 |
| o | TGAGCGCATC | TTTCATACGT | TCTTTATTTG | TACGATAAAT | ACGATATTCT | TTTGTTACAT | 2940 |
| | TTTTAATGTT | TACCGAAACG | TTCATTTGTA | GACCTTCCTT | ATTCACATTT | ATCTAGATTA | 3000 |
| | TAATATACTA | CTCAACAGTT | GTTAAATTTT | AAAACCTGTT | GTAAAGTGTA | TAGAAGATTT | 3060 |
| 5 | TGTTATTATC | AGAGTGGGTG | TTTTGACACA | AAATGTTAAT | CATCAATGAT | AACAATGATA | 3120 |
| | TTTAAAAACT | AAACTTATTT | CAACTTACAT | GATTGTATAC | TATAATGTAT | TTGTAATAAA | 3180 |
| | CTAATATTTT | AAAGAACTAG | ACAATAATTT | TGATAGCATC | CATGTATAGT | GATAGTATTT | 3240 |
| 0 | ACAACAATTA | TTATAATACT | ATTTAGTTAA | GTAGAGAAAT | AGTTAAACAT | TTGAAAGTGT | 3300 |
| | GGTTTAATGG | AATGTCAGCA | ATAGGAACAG | TTTTTAAAGA | ACATGTAAAG | AACTTTTATT | 3360 |
| _ | TAATTCAAAG | ACTGGCTCAG | TTTCAAGTTA | AAATTATCAA | TCATAGTAAC | TATTTAGGTG | 3420 |
| 5 | TGGCTTGGGA | ATTAATTAAC | CCTGTTATGC | AAATTATGGT | TTACTGGATG | GTTTTTGGAT | 3480 |
| | TAGGAATAAG | AAGTAATGCA | CCAATTCATG | GTGTACCTTT | TGTTTATTGG | TTATTGGTTG | 3540 |
| 0 | GTATCAGTAT | GTGGTTCTTC | ATCAACCAAG | GTATTTTAGA | AGGTACTAAA | GCAATTACAC | 3600 |
| | AAAAGTTTAA | TCAAGTATCG | AAAATGAAcT | TCCCGTTATC | GATALACCGA | CATATATTGT | 3660 |
| | (2) INFORMA | ATION FOR SE | Q ID NO: 13 | 73: | | | |

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13868 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

50

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 173:

| | ATTAATCACT | TGTTGTGTAG | AGTCTTGTCC | GTTTTGGTTA | TGATTGTTAG | CCATGATATA | 120 |
|----|------------|------------|------------|------------|------------|------------|------|
| | CCTCCCTTAC | AACACTCGTG | GACCAGAAGT | TTTCTGATCT | CTCACATTAA | CTTCTAACTT | 180 |
| 5 | ACGTACTGGC | ATTTCTGTGA | AATATTCTAC | ATTCTTTTTA | ATATCCGAAC | GAATTGCTTC | 240 |
| | AGTTAAAGAT | TGAACTTGAA | CATTATTTGG | TACGAAAAAG | TCAGTTTTAA | TGTCGATATA | 300 |
| | AGATTTATTT | TTTTTGTTAT | ATAGTTTCGC | AACTACATTT | GGTTGTCTTA | CTTGATCATA | 360 |
| 10 | TTTTGCAACC | GTATCGAATG | CCGTCTTTTC | AACAGCTTTA | CGAGATACGT | AAACATGACC | 420 |
| | ATCATCGAAG | TCTTTGTATA | ATCCAGGTTT | TCGATGCGTA | GGTTTGAAGA | TACTAAATAC | 480 |
| 15 | TAATATAAGA | CCTATTAATA | TCAATAGTGC | AGCAAGTGAA | ATAAGTAATG | GTTGGAACCA | 540 |
| | TTCAAATTGA | AGGAAGTAGT | CTTGATATTC | AGTTATACGT | CCATCTTGGA | TATACATGAA | 600 |
| | TAACAGGAAC | CCCACGATTA | CTACTATTAA | TAAGCCAAGG | ATAAAGTTTT | TAAGTCGTTT | 660 |
| 20 | CACCCCTAAC | GACACCTCCT | TAGTTAAAGT | TAATTTAAAA | ACATATTAAA | TATGTACCCA | 720 |
| | TCAGTTTTTT | TCTTAAACAT | AATAAATTAA | TAACTTTAAA | TTTATTTTTA | ATATATAAGA | 780 |
| | TGAAGTACCA | TTTAGTAATA | TATTCCCTAG | TTTTTGTAAA | TAAAACCTCA | TTATTAATTA | 840 |
| 25 | ATTYTCGTCA | ATATGTTTTG | AAGAACGATA | TTCTAAAATA | TCTGGGTCAC | GATGTTTAAT | 900 |
| | TAAAACCTTA | TTACTATTTC | TCGGTTTCTC | CTCACTCAAA | GATTTTATAA | GCGACCATAT | 960 |
| | CATCGCTATA | ATGACCACGG | AAAATGGTAA | CGCAGCAATG | ATTAATAAAT | TTTGAATTGC | 1020 |
| 30 | TTGAGTACCA | CCTGTGTAAA | TCATGATGAT | TGCAAATAAT | GCCATAATGA | TACCCCAACT | 1080 |
| | CACTTTGACA | AATGACTTCG | GATTAATATC | ACCACTTGAA | CTCAACATAC | CTAAAACATA | 1140 |
| | AGTTGCCGAA | TCCGCTGATG | TAACAAAGAA | AATCATAATA | ACAAGTAAAG | TAATTAAGCT | 1200 |
| 35 | TAATACAAAA | CCTAGCGGAT | AATGTTGTAG | CGTCGCAAAA | GTTGCTGTTT | CTGTCGCAGC | 1260 |
| | TTTĄĞCAATA | TCGGCAATAT | GATTATCTTG | TAAGTAAATT | GCTGACGCGC | CGAATACCGC | 1320 |
| 10 | AAAGAATATA | AAGCAAACTA | ACGCCGGGAC | AAAAAGTACA | CCTAGAATAA | ATTCTTTAAT | 1380 |
| | CGTACGTCCT | TTTGACACAC | GTGCAATAAA | TATACCTACA | AATGGTGCCC | AAGATATCCA | 1440 |
| | CCATGCCCÃG | TAAAAGATTG | TCCAATTTTG | TAACCATTGG | AATTTTTGAC | CACCTGTCGG | 1500 |
| 15 | AATGCGTAAA | CTCATACTAA | AGAAATTTGC | AATATAATTA | CCTAGACCAT | TCGTAAATGT | 1560 |
| | ATTTAAAATG | TATAGCGTTG | GCCCAACAAT | AAAAAGACCA | ATAAGTACTA | CAAAAGCAAG | 1620 |
| | TAACATGTTG | ATATTACTCA | ACGTTTTGAT | ACCTTTATCG | ATACCTGACC | ATGCTGACCA | 1680 |
| 50 | AGTAAATAAT | ATGGTTGCAA | TGACAATCAA | GATTACTTGC | ATCGTGAAGT | TACTCGGTAC | 1740 |
| | АТТАААТАА | AAATGTAAAC | CTTCGTTTAT | TTGCAATGCA | CCGAAACCTA | ATGTTGCAGC | 1800 |

| | CATTGCCTTT | TCACCTAATA | AAGGCGTCAA | TGTAGCGCTG | ACTAAGCCAG | GATATCCTTT | 1920 |
|----|------------|------------|---------------|------------|------------|------------|------|
| | ATGAAAGCTA | AAATATGCAA | ACACTAGCGC | GACAATACCA | TAGACTGCCC | ATGCATGAAT | 1980 |
| 5 | CCCCAATGG | AAAAATGAAA | ACTGCATTGC | ATCATTAATT | GCAGATTGCG | TGCCAGCTTT | 2040 |
| | ATGAATAGGC | GTTAATTTGA | AGGCATGACT | GATTGGTTCT | GCCGTTGTCC | AGAACACAAG | 2100 |
| | TCCTATTCCC | ATACCAGCAC | TAAATAACAT | AGCAAACCAA | GACGGCAATG | AGAATTCAGG | 2160 |
| 10 | ATCTTCGCCT | TCTTCACCTA | ATGTAATGTT | TGCGTATCTC | GAAAATAAAA | TATACACACA | 2220 |
| | GACAAATAAA | АТĄАСТАААА | CGAGCAATAA | ATAATACCAA | GAAAAATGTA | GCGCAATAAA | 2280 |
| 15 | TGTAGTAATG | TTTTGCGTGA | GTTTTTCTAA | CTGTTTCGGA | AATATTGCTC | CAAAAGCAAC | 2340 |
| | AAATATCGTA | CATATCACTA | AAGATACCCA | AAACACTAGA | CTTACTGATT | TATTTTTCAT | 2400 |
| | AAATACAAAC | CCTTTCTGTG | TAATGGTAAG | TTCATACCCA | TAACTGCAAC | ATTTTAATCA | 2460 |
| 20 | TTTGTAATTT | TATATAGACA | CAATTAATAA | TGCCTCATCT | TTTAAAAATG | ATATATAAAA | 2520 |
| | CACACTCAAA | TTATTTATCA | TTGAGCAACA | AAGTATTTTA | TTGTATTTAA | GTAATGCCTT | 2580 |
| | TCTAGTGCAT | TATTGATTTG | ATACCTGCAA | AGTTGCCATA | TTTCCGTTTA | GAATCAATAG | 2640 |
| 25 | TCGCTAGACA | CAAAAAATAA | GTCGCCTATA | CAGTATTTTC | TGCATAAGGC | GACTITACTT | 2700 |
| | ACTAATCTAT | ATATTAATTA | CTAATTTTCC | AATCATTGAT | TGTTTTTCCA | ACAATTGATG | 2760 |
| | TGCTTGATAT | AAGTTTTCAG | GTGATAAACC | TTCAAAAACT | TGTGTCGTTG | TTGGTTGGTA | 2820 |
| 30 | ATGCCCTGAT | TCTATATTTT | TCGTAATATC | TTCTAAATAC | TCATGTTGTT | TAATCATATC | 2880 |
| | AGGCGTTCGA | TGAATTGGAC | GCGCAAACAT | AAATTCATGT | GTAAATGTTA | TACTTTTTAA | 2940 |
| | TTTTAATGCA | TTTAAATCTT | GATCTTCATT | AAAAGCTACG | ATAGTCGTAA | TATGCCCTAA | 3000 |
| 35 | TGGTTTTATC | AGTTCAATCA | TAGTATTGTA | ATACAAGTCT | GTATTATAGG | TGCAAAATAT | 3060 |
| | ATAATCTACT | AATGGAATTT | CTTTAAATTG | ACGCACTAAA | TCCTCTTTAT | GATTCAATAC | 3120 |
| 40 | GATATCTGCG | CCCATCTTTT | CACACCACTC | TGTTGTTTCT | TGTCGTGATG | CTGTTGTAAT | 3180 |
| | GACAGTTAAA | CCATACCGTT | TAGCAATTTG | AGTGGCTATA | CTGCCTACAC | CACCGGCACC | 3240 |
| | ATTAATGATT | AAGACAGACT | TCCCTTCGTT | TTCAGCAGGA | TTCGTAGAAA | TTTTAAATGT | 3300 |
| 45 | ATCAAAAAAC | GTTTCATATG | CCGTAATACC | AGTTAGCGGT | AGACTAACCG | CTTCATTAGC | 3360 |
| | ACTTATGTTG | TGTGGTGCTT | TTGCAACTAT | AGCTTCTGAC | ACCAATTGAT | ATGTCGCATT | 342 |
| | TGATCCTTGT | CTATTTGGCG | ATCCAGCATA | AAATACAACG | TCACCCGGAC | TAAATAATGT | 3486 |
| 50 | AACGTCTGGT | CCGATAGCTT | CAACAGTACC | AATAGCATCA | AACCCAAGTA | CACGAGGTGC | 354 |
| | ттелетелет | TOCATTICTO | CALC CALABOTA | ATCTACACCA | TTTACACTAA | TCCTATTTAC | 360 |

| | ATTTCCTTCT | TCCAATTTAA | AGGGCTTCTC | AAATCCTATC | ATTTTCATAT | CGTTTCACCT | 3720 |
|----|------------|------------|------------|------------|------------|------------|------|
| | CATTTATGAA | CTTATTTCTT | ATTATACAAA | ATAGAAGCCA | TGTGTGCTTA | TATCGCAGCA | 378 |
| 5 | TCATGACTCC | TTTTTCATTT | GAATATATAA | ATAATTACAG | ACGACTITCG | TATTAAATTT | 3840 |
| | TAGACTTATT | TCTACCATGT | TGCTGAACAA | ATTTACTTTA | GATAAAAAAT | TATTAAATTT | 3900 |
| 10 | TGGTCAATTA | ACAAAGTTAG | TTTGTTAAAA | CGTgATACTT | TATTATTCCG | TTACTTTAAT | 3960 |
| 10 | AACTTGTTTA | CCAAAGTTAT | CGCCAGTaAA | TAAATTTTTA | AATGCATGTG | GCGCATTTTC | 4020 |
| | AAAACCATCT | TCAATGGTTA | CTTGTGACTG | AATTTTACCT | TCTTGAACCC | ATGTTGCAAG | 4080 |
| 15 | CTGTTCACTA | GCTTCTTTAA | AAGCATTAGC | GAATTCACTT | ACCAAGAAGC | CTCTCATCAT | 4140 |
| | TACTTGCTTC | TTAATAAGCG | TACCTTGAAT | ACGTGGTCCG | ATATCGGCTT | CAGGATGATT | 4200 |
| | ATATGACGAA | ATTGCGCCAC | ATACTGGTAC | ACGTGCAAAA | CGATTTAAAT | GCTTAAATAC | 4260 |
| 20 | TTCATCGCCA | ACTGTTCCAC | CAACATTTTC | AAAATAAACA | TCAATACCAT | CTGGTACTGC | 4320 |
| | TTGTGCTAAC | GCTTCTGCAA | AATCCTCTTT | CTTATAATCA | ATACCAGCGT | CAAAGCCCAG | 4380 |
| | TGTCTCTGTT | AAATAATTTA | CTTTTTTGTC | GCCACCCGCA | ATACCTACTA | CACGGCAACC | 4440 |
| 25 | TTTAATCTTA | GCAATTTGAC | CTACAACTGA | ACCTACAGCA | CCAGATGCAG | CTGAAACCAC | 4500 |
| | AACAGTATCA | CCGGCTTTAG | GTTGTCCAAT | ATCAAGCAGA | CCATGATATG | CTGTTTGTCC | 4560 |
| | TGGCATTCCT | AAAACACTTA | AATATAAATC | AAGTGGTACA | TCTGTCGTTG | GAACTTTAGT | 4620 |
| 30 | AATTTGATCC | GCTTGGACAT | GATTAATGAT | TCGCCAAGGC | AACATACCTA | CAACGACATC | 4680 |
| | TCCTTTTTTA | TAATCTGCGA | GTGTCGAATC | AATTACTTTT | GCAACGACAT | GGCTAACAAT | 4740 |
| 35 | CGGTTTACCA | ATTTCAAAAG | GCTGTACATA | CGAATCTGCC | TTAGTCATAC | GTCCTCTCAT | 4800 |
| 55 | ATATGGATCC | ACTGAAATAT | ACAGCGTTTG | TACAAGTACA | CCATCGCTCT | CAAGTTTaGG | 4860 |
| | CGTGFCAATC | TCTTCaATTT | TGAATGTATC | CTCTTGAGGC | ATGCCkTCTG | GTATTTTGTT | 4920 |
| 40 | AAAAAGAATT | TGTTTATTTT | GCATCATTAA | TCACCTTTCT | TTATTTGAAA | CTTTTACTTA | 4980 |
| - | TTTGTTACTT | AAGCGTTAAG | TTTGAATTGT | GTCLTCGTGA | TGTCTGTATG | CAAATACATT | 5040 |
| | CTTAGTTGTT | ATATTTTGAC | TTAAGCACTG | ATTCATTCAT | GTAACTTCAA | CCACATTATA | 5100 |
| 45 | TTTGCTATAA | TCATAAATTT | AAAATGTTAC | GACTTAGACA | TTTTATGGAA | ACTCTCAAAC | 5160 |
| | AATAGATAAT | TTTTGAAAAG | CTCTAATATT | ACAAGCTTTT | TTGCCCCAGA | AAAACTAGCA | 5220 |
| | GTTGCTTTAT | TTTTTCCATA | AGAAGTCGAT | TAACTCATTA | GCAACATTTT | CATTCTCATG | 5280 |
| 50 | AAGCTGACTA | TGTTGTGCAG | GCTCACCTTC | ATATTTAGAT | TCTCGATAAC | TTTTCGGACT | 5340 |
| | ATTTCCCAGT | AAATATTTTA | ATGATTTCGA | AGAACTATTA | GACACTCTGC | CGTCTGAATG | 5400 |

| | ATCTTTTAGC | ACGCGTAATT | GCTGATAAGG | TTGATTCATT | CGACTTGGTT | TACCATCTTT | 5520 |
|----|------------|------------|------------|---------------------|------------|------------|------|
| | ATCAACTGTA | ATTTCATTGA | CATCTTCATT | CATATTTAAA | ACACCATTAA | ATGTCCCTGC | 5580 |
| 5 | AATATTCACT | TGTTTGTTTA | ACTGTGGCAG | TGACTTGTCG | TTACCATATG | TCATCATATA | 5640 |
| | TTGTGCAAAT | GTTAAGTTCC | CCATTGAGTG | ACCGACAAAG | TTGAATTTAT | CGAAATTGTA | 5700 |
| | TTCAGATTGT | AACTTAGTCA | GTACATTTTT | AAACCACGCA | GCATTCTTAT | CCAAATAGCC | 5760 |
| 10 | TTGTCTGTTA | TTTTCAAGTT | CAATTTTCAC | AATAGGATTC | ACTGCATCTT | TTCTTAGTTT | 5820 |
| | CCCTTTAAAT | GTCACTGCAC | CATCCTTTGA | AACGTAAGCA | GTGATGATAT | CTTTAGTTAC | 5880 |
| 15 | CCCTCTTTTT | TCTGCTTGCT | TCACCATAAA | CTTTTCAGAA | TTGGCACTAC | CACCAAATCC | 5940 |
| 15 | ATGTAAGAAC | AATGTTGGAA | TIGGCTITTT | AACAAATTGC | TGTTGTTGTA | TTTTAAATGT | 6000 |
| | TTGTGCCTGT | CGTTGACTAA | ACACCACCAT | AATAATAGAG | CCTATAATAA | TAGCGACCGC | 6060 |
| 20 | TAACAATGTC | GTAATAATTA | CAAAAATTTT | CTTCACACTT | TTAACTCCCA | TTCATGTCTT | 6120 |
| | TTATATAAGT | ATAAAGGATG | TGATTAAAAA | TGTCCTTTAG | TTGATTTTGA | ATACATCATT | 6180 |
| | AACTTTTAAG | ATGACTTTGG | AAAGTTGTCC | GTTAACGTTT | GTTAATTGAT | TGCTTCTTTA | 6240 |
| 25 | GCTTTCAATG | GTGTGTCACC | CATTGATTAA | TATATAAATA | TGTATATGCA | TGTTTAATTT | 6300 |
| | ATCTCTTTCT | ATAAATAAAG | ACCTACCAGC | ACTCGACTGA | TAGGCCTTTT | AATATCTATA | 6360 |
| | TAATTTATTA | TTCTTTTGTT | TCGGCTAACT | CTTTGTACCA | ATAAGCACTT | TTCTTAGGAT | 6420 |
| 30 | AACGTTCTTG | AGTCTCAAAA | TCGACATAGA | ATAAACCATA | TCGTTTTTCA | TAACCATTTG | 6480 |
| | ACCAAGAGAA | CACATCCATT | AATGACCAAA | TAAAGTAACC | TTTAACATTT | GCACCATCTA | 6540 |
| | TAATAGCATC | TGCAATAACG | TTCAAATGTT | GTCTTACATA | ATCAATACGT | GCATCATCAT | 6600 |
| 35 | GAACTGTTTT | TTCAGATTCA | ATAAATTCAT | CTTTATATCC | TAAACCATTT | TCAGTGATAT | 6660 |
| | AAATCTTATG | aTAGTTAGGA | TAATCTTTAA | CAACACGCAT | GaTTTGATCA | TATAAACCTT | 6720 |
| 40 | GAGGATAGAT | CATCCAGTCC | CAGTCTGTGC | GAGGTACGTC | GACATCAAAT | TCACGTTGTC | 6780 |
| 40 | CGACACCTTT | AAGTTGGTAT | TTAGAACCGC | CTTTATCACC | TGTCGCATTA | TGCGTGATTT | 6840 |
| | CAGATTCTCC | ATCGTAACCT | CTCATCCAAT | CACTCATGTA | GTAATTGATA | CCTAAGAAGT | 6900 |
| 45 | CGTTTAAATC | TTTGGCTGCA | TCTAAAATGG | CATAATCTTC | ATCTGTAATG | TTTAATTTAC | 6960 |
| | CGCCATTAAC | AGATAAGATA | TGTTGCACAC | CTTCCATCGT | TTCACGAGAA | TACTTACCTA | 7020 |
| | AATATGTTGC | ATCTAAGATG | TATTTATTAA | GGATGATATC | TTCTAATTCT | GCTGCACGAA | 7080 |
| 50 | CATCTTCAGG | ATTTGATGGA | TCGAACGGAT | ATTTTGTTGG | CAATGCGTGT | ACAACACCAA | 7140 |
| | | CTATCCCCA | TOTAL SATS | Variation Variation | тстассатса | CCCACCATCA | 7200 |

| | СТАСТАААТА | TTGACCATCA | CCAATAGGTC | CAATTTCATT | GAATGTAGTC | CAATATTTTA | 7320 |
|----|------------|------------|------------|------------|------------|------------|------|
| | CTTCTGGGAA | TTCTTTAAAA | CAATATTCAG | CATAATCTAC | AAAGTAGTCA | ATCGTTTTAC | 7380 |
| 5 | GATTTAGAAA | ATCGCCATCT | TTGTGTAACA | CTTCTGGTGT | ATCAAAATGA | TGCAATGTTA | 7440 |
| | CAAATGGTTC | AACATGACGT | TTATGACACT | CTGCAAATAA | CTTATGGTAA | TACTCAACAC | 7500 |
| 10 | CTTTAGGGTT | AACTTCGCCA | TATCCATTTG | GGAAGATACG | AGACCATGCA | ATTGAAATTC | 7560 |
| 70 | GGATACCATT | AACACCGAAT | TTTTCACTTA | ATTCTAAATC | CACTGGATAT | CTGTTATAAA | 7620 |
| | AATCACTCGC | TGGTTCTGCA | GTGTACCAAT | AGTTTTCTTC | TAAATACGTA | TCCCATGCTA | 7680 |
| 15 | CGCGACCTTT | ACCATCCGTA | TTTGTCGCAC | CTTCTGCTTG | ATATGCTGCT | GTTGCTCCAC | 7740 |
| | CAAAAATAAA | ATCTTCAGGT | AATGTTTTAG | TCATATGAAA | AACTCCTATT | CTTAATTTTC | 7800 |
| | AAATTGTTGT | TGAACGAAAT | CAAGGGCTGC | TTGGCCATCT | CGTGTCAATT | TGATATATTC | 7860 |
| 20 | AGCACCTTGA | GTCTTCGCTA | ATTTAATACC | TAATCTATCT | GTATCTTGCT | TAATATCTTC | 7920 |
| | ATAGTTAGAC | GCAACTTGTG | GCGCTAAAAT | GATTAATTGG | TACTCTTTCA | TAATGTCCAT | 7980 |
| | ATGTGCGCCA | TATCCGCCAG | cTGCCGCTTT | CACTGGCACA | TGATATTCTT | CAGCTGCTTT | 8040 |
| 25 | ATTAAGTGCA | TTGGCTAATA | ATCCACTTGT | ACCACCACCG | GCACAAAGTA | CTAAGACATT | 8100 |
| | TGTTTGTTCT | GTGATATTTG | AAGCTTTAGC | TGCATCGTCT | GATACACCAC | TTGCCGCTAA | 8160 |
| | AATTGAATCA | GCTTTTTTCG | TATCAAAGTT | TGCTGCAACT | TTTTCTTTTA | AATCTGAATT | 8220 |
| 30 | ACTITCTITA | CGTCCTTCTT | CTTCATCAAG | AATTTCACTA | TCATAAACTT | TTAGGAATGG | 8280 |
| | GTAGTAAATA | ATAATATCTA | CAACAATCAA | AGTAATAGCT | AGTACGAATG | ACCATAAACC | 8340 |
| 35 | AAAACCTGTA | CCCATGATAA | TGCCCAATGG | ACCTGGTGTT | GTCCAAGGTA | AATTCACACT | 8400 |
| ,, | AAAACTATTC | ATTCCTAACA | CTTCAACGAA | AAGTTTGAAA | ATCCATACGT | TAACAATTGG | 8460 |
| | TGCTAATACA | AATGGAATAA | AGAACACAGG | ATTCAATACT | AGTGGTGCAC | CAAATAAAAT | 8520 |
| 10 | TGGTTCGTTT | ACACCAAAGA | ATGTTGGTAC | AACTGATGCA | CGTCCAATCG | CTTTGTTTCG | 8580 |
| | TTTAGATTTC | GTCATCCACA | TAAACATGAA | CGGGACGACC | AATGTTGCAC | CCGTACCTCC | 8640 |
| | AAATGTAACG | ATAAACATTT | GTGTACCTGA | TGTAATAATT | TTATCTGCGT | GTTCTCCAGC | 8700 |
| 15 | TTGAAGCAAC | TTGAAGTTCG | CTTCGATATT | CGCATATGTA | ATGGCTGCAA | TTGCTGGCTC | 8760 |
| | TACAATTGAC | GGACCATGAA | TACCTACAAA | CCAGAATAAT | GCAAAGGCAC | CAAAGATAAT | 8820 |
| | TGTGACACCA | ATCCATCCAT | стестестет | AAATAATGGT | TCGAATAATT | TTAAAATACC | 8880 |
| 50 | TTCCGCTACA | TTTGATTTAA | AGCTGTTGCG | AATGACTAAA | TCTAATGCAT | AAAGAATGAT | 8940 |
| | GATTACCGCT | GAAAATGGAA | TTAAGTCCTT | AAATACTTGT | GAAATATTCG | GCGGTACTTC | 9000 |

| | AAATGCTGAT | AAGAATGCTG | TTAATAAACC | TTTAGTTCCC | ATAAATGCAC | TTAAGAATCC | 9120 |
|----|---------------|------------|------------------|----------------|-------------------------|--------------|-------|
| | ACCATCTTTG | GCTGGATCAG | AAGCTAAGAA | CAAGAATCCA | CACATCGCTG | CTAGCATTGT | 9180 |
| 5 | AGAAATAAAG | TTAATTTGAT | TTGTACTTTC | TAGCTTACGG | TTAAATGAAT | CTGTTAAAGA | 9240 |
| | TTTCGCTGTC | GTTCCTGCTA | CTAAAAATGC | TACAAGCCCC | ATCGTATAGT | TATATGGTTT | 9300 |
| | CATTAAAATG | GCTTCCATGC | CTTTATCCCA | TTTAAAACCA | AATATATTTG | GCACATATGC | 9360 |
| 10 | AATTAATAGA | AAGATACTTG | AGAATAAGAT | GACAGGCATT | GCAGAAATAA | ATCCATCACG | 9420 |
| | GATGGCTCTT | AAATATATGT | TACGTGATAA | TTTCTCGAAA | AATGGCTTCC | CTTTTTCAAT | 9480 |
| 15 | TTGTGCGATC | AATTTTTGCA | TCATTGTCAT | CACCCTCTTT | TATAAAATTC | TAATAAATGC | 9540 |
| 13 | TTCATTAAAT | CTTTCAGTAA | AATTGTTGTC | ATTAAATGAT | CTTGACCATG | CATCATCGTT | 9600 |
| | ACACTGTATG | CAATATCATC | ACCTTGCGCT | TCTTTAGCCA | ACAGGCTTGT | TTGTGCTCTA | 9660 |
| 20 | TGCGCTTCCG | CAATGCAATT | GTTTCCTTCT | TCAATCAGTG | CATCTGCTTT | TGCAAAATCT | 9720 |
| • | CCAGCTTGAG | CTGCTGTTAA | TGCTTCTAAA | AACTTAGAAC | GTGCATCCCC | TGCAAATGCA | 9780 |
| | ACAATTTCAA | AACCTAATAA | TTGGACTTCT | TCTCTATTCA | TAGCATTAAT | CCCCTTTTAA | 9840 |
| ?5 | ACTTATTTTC | TTTGTTTCCA | AGATGTCGCA | GTATCTTTTA | ATACTTTATT | TAAGTCATCA | 9900 |
| | ATATTTTTGA | AACCAGTTGT | ACGTAACCAT | TCACGAGCAG | CATCTTCACC | TTGTTCAATG | 9960 |
| | TATACTTGAA | CAGCACCAGA | CCATGTAGCA | CGGCCACAAA | GTACCCCGTT | GAATTTAGCA | 10020 |
| 30 | CCAGCTTCGT | GAGCAAATTT | TAAAGTTTCT | TGGAATAATT | CCGCAGAAAC | ACCAGCACTT | 10080 |
| | AAGTAAATGT | ATGGTAAATG | AGTTGCTGCA | TCTTGATCTT | TAAAGTGTTG | TGCCGCTTCC | 10140 |
| | TCTTTTGTAT | AAACCACTTC | ACCTTCAGCA | AAGCCTTCTA | CATATTTCAT | GTTTACTGGT | 10200 |
| 35 | ACTTCAACTT | TCAATACATC | AACGTTAAAG | CGTGGTTCTG | AGAATAATTT | CATTGCTTCG | 10260 |
| | TTAACCTTTC | TAGGCTTAAC | TTTTGCGAAT | TCAACACTAC | CGTTATCAGG | AATGTTGTCA | 10320 |
| 40 | TCGTATGTTA | ATACTTCTAA | AAAGAATGGA | ATATCTTCTG | CAACACATTC | TGAACCGATT | 10380 |
| •• | CTTTCAATGT | ATGCTTTCTT | TTGAATGTTA | ATTTCTTCAG | CATCATCAAC | ATCATAGTAA | 10440 |
| | AGTAAGAATT | TAACAGCATT | TGCGCCTTGT | TCTTTTAAAC | GTTTTGCAGA | CCACTCTACT | 10500 |
| 45 | AAACAGTCAG | GTAAACGACC | TTTAGCGTTT | ACGTCATATC | CAGTTTTTTC | ATAAGCAAGT | 10560 |
| | AATAATCCAC | AATCTTTGTT | ACGTGCATCT | GAAGCTGGTA | AACCATATTC | AGGATCTAAT | 10620 |
| | AAAATTGAAG | ATGCATATTG | TGTTAATTCT | TCCGCAACTA | ATACTTTTAA | TTGTTCAATT | 10680 |
| 50 | TGAGCTACAG | TTGGTTCTTC | AGTTTGATGT | TTTGCCATCA | TGCGTTTTAA | AGCACCACGT | 10740 |
| | TCCTCN N N TC | CTARTCCACA | 3 8 TC 8 T 8 CCT | TACCOTATION CO | THE STATE OF THE PERSON | A ATTICATION | 10800 |

| | TCATCATAA1 | TATTTAAATT | GACATAACCI | GTTTGTGCTT | CTTGTGCATT | CAGCATGCCT | 10920 |
|------------|------------|--------------|------------|------------|------------|------------|-------|
| | AAAGTATTGG | CTTTTTTAG | TAAATCGTGG | TCGTTTTCAT | GATTAAGAAT | TGCTGAAGTA | 10980 |
| 5 | ATTCCAGCAA | CTGTAGAATC | ACCTGAACCA | ACCGGATTTA | ATACACTTAT | TGTCGGAATA | 11040 |
| | TTCACTCTAT | ' AGAATGTATG | ATTGTGCTTA | GCGAATGCAC | CTTGTGCACC | TAAAGACACA | 11100 |
| | ATAATCCACT | CAATCCCTTC | GAATAAGGGT | TGTGACACTG | CCTGTTTCAA | ACTTTCTAAA | 11160 |
| 10 | CTTTCATCAA | GTGGCTGGTT | AAGCAATTGA | TATAGTTCAG | AAATGTTTGG | TTTAATGACT | 11220 |
| | GTAGGTTTGT | ATGGATTTTC | CAAAACTGTT | TGCAAAGTtG | CACCCGAGCA | ATCTAATATC | 11280 |
| 15 | ACAGGCACAC | CTTTGTTTTG | GCATCGTTCA | ATGATTTGTG | CATAATAATC | TTGATTTAAT | 11340 |
| | CCTTTAGGTA | AGCTACCTGA | AATAGCAACT | GCTTCAACTT | TTTCTAATAA | TTGTTCAAAA | 11400 |
| | TGTTTAATAA | ATCCTGCAGC | CTCTTGATTA | TCAATCTCCG | GTCCCTGCTC | TAAAATTTCT | 11460 |
| 20 | GTTTGTTGCC | CTTCATGTAA | AATTGCAATG | CAGTTTCGTG | TTTCACCCTT | AATGTTATAA | 11520 |
| | AATGCATGCT | TGATGTCGGC | ATGATCTAAT | TTTTTAGCAA | TAAATTGACC | TAATTCACCG | 11580 |
| | CCAATAAAAC | CACTCGCAAG | GACTGGCTCA | CCTACTTGCG | CAAGTACTCT | TGTTACATTT | 11640 |
| 25 | AAACCTTTAC | CACCAGCTGT | TTTACTTACT | TCTTGAACAC | GATTAACATC | ATCTAATTTC | 11700 |
| | AATGCTGTTA | ATGGGTATGA | AATATCAACG | GATGGATTTA | ATGTTAAAGT | TAAAATCATA | 11760 |
| | TGTGTCGTCC | CTTAATCGTG | GTATTCGCCT | CTGTCCCATT | TTTCTAAGAA | TTCATCAAAG | 11820 |
| 30 | AAATGTGGAT | CAGCTTGATC | TGCATTGCTT | GTTTCTAAAT | GTTTAATTTT | AGCGATTAAT | 11880 |
| | TTTTTGTTCT | CTTCAGTTGG | TTTATATTCA | GCATTAATAA | ATGCATCGAT | AATATCGCAC | 11940 |
| 35 | ATTAATAACT | CACCTATAAT | ACGTCCACCG | AAGCCAATAA | CGTTCGCATT | TAATTCTTCT | 12000 |
| 33 | TTAGCGTATA | ACGCTGATGT | CATATCACGT | ACTAGTGCTG | AACGAACGCC | AGGTACTTTA | 12060 |
| | TTTACAGCAT | TGTTAATACC | AACACCTGTT | CCACAAATAC | AAACACCTAA | GTCTGCATTA | 12120 |
| ‡ 0 | CCGCTAACAA | CTTGTTCGCC | AACTTTTTTA | CCAAAAATTG | GATAATGTGT | TCTTGTGAAA | 12180 |
| | TCGTATGTTC | CTACGTCAAT | GACTTCATGT | CCTTTTGATT | TTAAAAATTC | AGATACACGC | 12240 |
| | ATTTTTGTAT | CTGTAACAAT | ATGGTCGCAT | CCTAATGCAA | TCTTCATAGT | AATTTTTCCT | 12300 |
| 15 | CCTTAGCACA | TTTTATTAAG | CATATCTACG | CGGATTTGGT | GTCTACCACC | ATCGTATTTA | 12360 |
| | CCTTCAACAA | AACCTTTAAC | GACATTTTTC | GCTAATGTGT | CTCCAACAAT | TTCAGATCCC | 12420 |
| | ATAGTGATCA | TTCTTGAATT | GTTATGGCCT | CTAGTCATAT | ATCCAGAGCG | TTCATCTGAT | 12480 |
| 50 | ACTTCAGCAG | CAATCATGCC | TTTGATTTTT | GTAGCAACCA | TAAAGCTACC | TGCACCAAAT | 12540 |
| | GCATCGATAA | CAATACCTAA | GTTACCTTCT | TGACTTTGAA | CATCTTTTGC | TACACCCAAA | 12600 |

| | TCTAATAAGT | ATGATTTGAT | GACTTCTTTT | AATCGTTTGC | CAGCTTCATC | TGAACCAATA | 12720 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | ATAATCGCCA | TAATAAGACT | CCTTTTTACT | TTAATTTTGA | AATACCTTTC | TTAAAATGTG | 12780 |
| 5 | ACATATTTAT | TTGTAGGTTA | TGAAAATCTT | GAGAAAAGGC | TTTCAATTTG | ATTACGTTTA | 12840 |
| | AATTATAAAC | ATAAACAAAC | AATAAATCAA | CATAATATGT | TTATAATATG | TTTGTTTATG | 12900 |
| | ACGTATTTTC | AAACAATAAG | TGAACATTCA | TATTGTGGTG | TTGTTTTAAT | TAGGTATTCG | 12960 |
| 10 | TCTGAAATTG | TAGTAAAACT | TTGTCGAGGT | TCCCGTTGaC | ATAAATTTGC | ATAAAAAAtA | 13020 |
| | GCCCATAAAT | GAATGCAAAT | TCACATTCAC | TTATGAGCAT | ATAGATACAT | ATTTTAACAA | 13080 |
| 15 | TGCAGTTATA | CTTTTAATTT | AGTCGACTAC | TTCAATATAT | GTTTTAATCG | TTTCTACTTT | 13140 |
| | TTCTTCATCT | TCATAGTCCA | TGACCACTGC | AGTCAATTCG | TTTAACTGAC | AAAATGATGT | 13200 |
| | AAAATCTTCT | TTGCCAACTT | TCGTATGATC | GATTAACAAG | TATTTTTCAA | TTGAATTACT | 13260 |
| 20 | TAGTGCCAGT | TGTTGCGTAT | AGGCTTCATC | TAATGTAGAT | GTCATCACAG | CACCTTTATT | 13320 |
| | TACTGCGTTA | CTACTAAAGA | ACATCTTGCT | AAATCTTAGT | TTTTCCAACA | TGGCGTTCGC | 13380 |
| | CATTTCACCT | ACAAATGCTT | CTGTAATATG | GCGCATTTCA | CCACCAATTA | AATAGACACG | 13440 |
| 25 | AAAATGTGCT | GTTTGTTTTT | CTAACAAAAT | TTTATACACC | GGCAAACAAT | TCGTAATAAT | 13500 |
| | TGTGAGCGTA | TGATGATTGA | CTTCTTCTGC | TAATAGTTCC | ACTGTTGTTC | CTGGTCCGAA | 13560 |
| | AAACAAAGTA | TCCCCATCTT | CAATTAATGA | TGCAGCTTTT | TTAGCTATAA | ATCGTTTTTC | 13620 |
| 30 | TGCAATTTGA | CGGGTATGTT | TTTCTTTATG | CGATATTTCT | TTATACTGAA | ATGTTGAATT | 13680 |
| | ACTGCGTGCA | CCACCATGAA | TCTTCGTTAA | AATCCCTTTA | TTTTCCAATT | CAATTAAATC | 13740 |
| | TCTTCGAACT | GTCATATCAG | ACACATTTAA | ACCTTCGACG | ATTTCATTCG | TTCTTATCGT | 13800 |
| 35 | GCCCTTTTTA | TTCACTAGTT | TAGCAATTTC | GTCCAAACGT | TCATGTTTAT | TCAATGTAAA | 13860 |
| | ATTGCCTC | | | | | | 13868 |

(2) INFORMATION FOR SEQ ID NO: 174:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4549 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 174:

60 TTAAGTCAAC TTTGTCTATA CGGTTTGGAT CLTCTACCCA ATGTCTTATA AAAGACAATC 120 CCGCACCTGA AACATAACTC ATGAAATAAG AAAATGGTAT ACCATTAATT TGATCATTTT

55

40

45

| | AATCTTTACC | CATACGAAAC | ATCAATTGAT | AAAATGCGAT | GICTITITCI | ATCATTTCTA | 240 |
|----|------------|------------|------------|------------|------------|------------|------|
| | TTAAAACGGT | CATAATTTGA | TGTATGTTAT | CCGTGGATAA | CTTAACTGCT | CCATTTAACT | 300 |
| 5 | TCTCATCATG | AATGAAGTCT | CTTATTTCCT | CCAACTGCTG | GTCCTCTAAT | TTTTCAAGCA | 360 |
| | AATCATACTT | ATCATAATAA | TGCGTATAAA | ATGTACTACG | GTTAACATCA | GCTAAATCTG | 420 |
| 10 | CAATTTGTTG | CACAGTAATC | TCTTCTAATT | GGTGTTGATG | TAAAAGTTCA | ATAAATGCAT | 480 |
| | TTCTCATTGC | AACTTGTGAT | TTTCTAATAC | GTCGATCTAT | AGTCATTTAT | ATCAAGTCCT | 540 |
| | CCCCAATGAT | TATAAACGTT | ATGTTCATTA | TCCCACAAAT | CTCCAACATT | GATGATTGGC | 600 |
| 15 | ACACAATGTT | TACCTGTTTA | ATATAGGTGA | TACAAACAAA | CAGAAAAAGG | TGATAACAAT | 660 |
| | GAACCAACAT | TTACTAGGAA | ATCCAAAATT | AACTGTAACT | CATGTCAATG | AAGTTAAAGC | 720 |
| | CGGTATTAAC | CACATCGTTG | TCGACAGTGT | TCAATATGGA | AATCAAGAAA | TGATTATGGA | 780 |
| 20 | AAAAGATGTC | ACTGTGGAAA | TGCGCGATGG | CGAAAAATTA | TATATTAATA | TTTTCAGACC | 840 |
| | AAATAAAGAT | GGCAAATTCC | CTGTAGTTAT | GTCTGCAGAT | ACTTACGGTA | AAGATAATAA | 900 |
| | GCCTAAAATC | ACAAATATGG | GTGCCCTTTG | GCCAACATTA | GGTACCATTC | CGACATCTAG | 960 |
| 25 | TTTTACACCT | GAAGAATCAC | CAGACCCAGG | ATTTTGGGTG | CCAAATGATT | ATGTTGTAGT | 1020 |
| | TAAAGTTGCA | TTACGCGGTA | GTGACAAATC | CAAAGGCGTC | TTATCTCCAT | GGTCAAAAAG | 1080 |
| | AGAAGCGGAA | GATTATTACG | Artgattgaa | TGGGCAGCAA | ATCAGTCATG | ĠAGTAATGGA | 1140 |
| 30 | AATATCGGGA | CAAATGGTGT | TTCTTATCTT | GCGGTGACTC | AATGGTGGGT | CGCATCATTA | 1200 |
| | AATCCACCAC | ATTTAAAAgC | AAtGATTCCT | TGGGAAGGCT | TAAATGATAT | GTATAGAGAA | 1260 |
| 35 | GTAGCCTTTC | ACGGAGGTAT | mCCAGATACT | GGCTTTTATC | GTTTCTGGAC | TCAAGGTATT | 1320 |
| | TTTGCGAGAT | GGACAGATAA | TCCAAATATC | GAAGATTTGA | TTCAAGCACA | ACAAGAACAT | 1380 |
| | cctciettcg | ATGATTTTTG | GAAACAGCGT | CAAGTGCCAT | TATCACAAAT | TAAAACACCT | 1440 |
| 10 | CTACTAACAT | GTGCTAGTTG | GTCTACACAA | GGTTTGCACA | ACCGTGGCTC | TTTTGAAGGA | 1500 |
| | TTTAAACAAG | CTGCATCTGA | AGAAAAATGG | CTATATGTGC | ATGGACGTAA | AGAGTGGGAA | 1560 |
| | AGTTACTACG | CTAGAGAAAA | TCTCGAACGC | CAAAAATCAT | TCTTTGATTT | TTACCTTAAA | 1620 |
| 15 | GAAGAAAATA | ACGATTGGAA | AGATACGCCT | CATGTCATTT | ATGAAGTTAG | AGATCAATTT | 1680 |
| | TATAAAGGCG | AATTCAAATC | AGCGTCACGT | GTCCCTTTAC | CTAACGCAGA | ATATACACCA | 1740 |
| | TTGTATTTGA | ATGCTGAAAA | TCACACATTG | AATCATGCAA | AGATTAGTAG | CGCGCATGTC | 1800 |
| 50 | GCACAATATG | ACTCTGAAGA | TAAACAACAA | GATGTAAGTT | TTAAATATAC | GTTTGACAAA | 1860 |
| | GATACTGAGT | TAGTTGGAAA | CATGAACTTA | AAACTATGGG | TAAGCACTAA | AGACTCAGAT | 1920 |

| | CCTGATTTTA | ATCATATTGA | AAATGGTCAA | GTAGCTACTG | GTTGGTTACG | CGTATCACAT | 2040 |
|----|-----------------|--------------|--------------|------------|------------|------------|------|
| | CGTGAATTAG | ATCAAGAAAA | ATCCTCAATC | GCGCAACCTT | GGCATAAACA | TGAAACAGAA | 2100 |
| 5 | TTAAAGTTGT | CACAAGATGA | GATTGTACCT | GTTGAAATCG | AATTGTTACC | TTCAGGCACG | 2160 |
| | CTATTTAAAC | AAGGCGAAAC | ATTGGAAGTT | GTTGTAAAGG | GTAGTGAAAT | TGTAATTGGT | 2220 |
| | AATAGTACTC | CTGGCATGAA | AACACGTTAT | GAACATGAAG | AAACCGTAAA | TAAAGGCATG | 2280 |
| 10 | CACATGATTT | ATACTGGTGG | TAAATATGAT | TCACAATTAA | TCATTCCTAT | CGTTAATTGA | 2340 |
| | TATGCAGCAA | TTACGGTCGC | TTTTGATTAA | AAGTGACATA | GTGATAGGAC | TGTATAACAA | 2400 |
| | GAGAAAGCCA | CACGCTTGGA | ATCTTAAACC | AAGGTGTGGC | CCTTTTTATT | ATTGATGGCT | 2460 |
| 15 | ATTTAATTTT | ATAACACTAT | CGTATTTTCT | TTTTCATGAA | TCATTTCAAT | AATGACATTA | 2520 |
| | TCTTCATTCA | TTACTGCTAC | TTTAGGTGCA | TGGTTTTTAA | TITCTTCTTC | ATTCAACTGT | 2580 |
| 20 | GCATAAGTCA | TGATTATGAC | TACATCGCCT | ACTTCAACAA | GTCTTGACGC | TGCACCGTTT | 2640 |
| | AAACAAATTT | TACCACTACC | TCTTTCACCA | GCTATTACGT | ATGTTTCAAA | ACGTGCACCA | 2700 |
| | TTATTATTAT | TCACGATGGC | TACTTTTTCA | TTTGGCAAGA | TGTCTACCGC | TTCCAATATA | 2760 |
| 25 | TCTGAATCAA | TCGTAATGCT | ACCTACATAA | TTTAAATTTG | ACTCAGTCAĈ | TCTTGCTCTA | 2820 |
| | TGAATTTTAG | CATTCATCAT | TGTTCTTATC | ACTTTATTCA | GCTCCAATTA | TTATATTATC | 2880 |
| | TATTAAACGC | GCTTTTGAAA | ATTTAACAGC | TAACGAGATA | AATATGCGTC | CAGTTATTTC | 2940 |
| 30 | GTGTTGTTCT | ACTAATTGAG | GATAACTATA | AACAGCAACT | TCTTCAATGC | GTTCACTTAT | 3000 |
| | ATGTGATTCA | AGATATTCAG | TAACCCTGTC | TATAATTACT | TTACTTTGAC | GTTCACCGTC | 3060 |
| | TTGATACAAC | GCTTGTGCTA | ATAGCAAACT | TTTACTTAAA | TGTACCGCTT | CTTGTCGTTC | 3120 |
| 35 | TTGCTCCGTT | AAATAAACAT | TTCTTGAACT | TTTCGCCAAA | CCATCTGCTT | CTCGAACGAT | 3180 |
| | ATCAATACCA | ATAATTTCAA | CGGCATGATT | GAAGTCTTTT | ACCMTTTGCT | CGaCAATAGC | 3240 |
| | CAATTGCTGG | GCATCTTTTT | TACCAAAATA | AGCATAATCC | GGCATAACAA | TATTAAATAG | 3300 |
| 40 | CTTATTAACT | ACTGTTACCA | CCCCATCAAA | ATGCCCTGGr | CCGETCGCTC | CTTCTAACAC | 3360 |
| | ATCAGCTAAT | GGGCCTACTT | TGACATCAAT | ACCTAATTCA | CCTGGATACA | TATCTTCTAC | 3420 |
| 45 | TGCAGGATGA | AAAACAATGT | CCGCTCCTAC | TTCTGATACT | AATTCTAAAT | CTTTATCAAT | 348 |
| | TTGTCTCGGA | TAAGCATCGA | AATCTTCGTT | TGGACCAAAT | TGTAATGGAT | TAACAAATAC | 3540 |
| | ACTCACAATT | GTAATATCAT | TTGTACTAAC | TGATTCGCGT | ACCATCGTTA | AATGTCCATC | 360 |
| 50 | ATGTAAGGCA | CCCATTGTTG | GGATAAAACC | AATCGTTGTG | CCTGAGCGTT | TGGCTGCTTT | 366 |
| | A A C A A MCMCM | mcca memeran | THE COURT OF | ****** | CTCATTCTTA | ттаасстсат | 372 |

| | GATCGTATTG | TTTTAAACCA | TCCACACCAA | CACTAAAATC | AGCAAATTGC | TTCACAAATT | 3840 |
|----|-------------|-------------|------------|------------|------------|------------|------|
| _ | TCGCTTTATG | TTCAACACCA | TAATTTAACA | TATCGTGATA | AACCAATACT | TGACCATCTG | 3900 |
| 5 | TACCTTTTCC | TGCACCAATA | CCAATGACTG | GAATTGTTAA | GTGCTTGCTA | ATTTCTTCTG | 3960 |
| | CTAAATCATT | TGGAATTGCT | TCAAGTACTA | ACGCAACTGC | ACCAGCTTGT | TCTACATTTT | 4020 |
| 10 | TCGCGTCTAA | AATAAGTTGc | TCCGCTGCTT | CTTTCGTTGC | ACCTTGTAAT | TTATACCCCA | 4080 |
| | TAACGCCAAC | ACTTTGAGGT | GTTAATCCTA | AATGTGCAAC | AACAGGAATA | CCAATTGCCG | 4140 |
| | TTGCTTTTTC | AATAAATGGT | GTAATATGCG | CTCCTTCTGC | TTTAATTGCA | TTTGCATTCG | 4200 |
| 5 | TCTCCTGATA | AAGCTTTAGA | GCATGATTTA | AGTCTTGTGT | CATAGAGATG | CCTACTGCAC | 4260 |
| | CAATCGGCAT | ATCAACAACT | ACAAATGTAT | TTGGTGCGCC | TCTTCTTACT | GCACGACCGT | 4320 |
| | GATGAATCAT | ATCTGCTAAC | GTCACTTGTA | CGGTACTTTC | ATAACCTAAT | ACAGTCATAC | 4380 |
| o | CAAGTGAATC | CCCAACAAGA | ATCATATCAA | TACCCGCTGC | TTCCACTTGT | TTAGCACTTG | 4440 |
| | GAAAATCATA | AGCTGTTACC | ATAGAAATTT | TAGTTTGCTT | TTGTTTCATA | TCTATTAATT | 4500 |
| | GACTTACTGT | TTTCAATGTT | ATTCAACCTC | TTTTTGCAGT | ATnATTAGA | | 4549 |
| 5 | (2) INFORMA | TION FOR CR | 0 70 110 | _ | | • | |

(2) INFORMATION FOR SEQ ID NO: 175:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8339 base pairs (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 175:

| TTATCTTTTG | TTGTTTCCTT | AGACAAACGA | CTAACCACAT | TATAATGGAC | TAATTTATTA | 60 |
|------------|------------|------------|------------|------------|-------------|-----|
| ATTĒTATTTA | ATTCCATTAA | GTTATCCGTA | ACACTAAGTG | AAGATGCGGA | GTTCACTCTC | 120 |
| GTTTGTACTC | TTCGTTTTAA | TAAAGCACCT | CGTAATAATA | CAATCATTCT | TCTTATTAAT | 180 |
| GATGCTTGTC | TATATACCTG | TGTTCTTTCA | GCATAACGCA | TATAGTTTTC | AAGTACACTA | 240 |
| TTCGTTATTT | GTCCTTCATC | TACTAAAGAC | TCTAATGTTT | TKGTTTCTAC | ATTAAAAGCA: | 300 |
| ATTTTTTGTA | GACGTTCTAA | TTCTTTAGAG | TTTTCATCAT | CTTTCTCTAC | AGTTTTTAAA | 360 |
| AATGCTAATT | TATCATGATA | TTCTTTAATC | ACGTTACCAT | ATTTAAAACT | TGTTTCGAAA | 420 |
| GTAGATTTTT | GATTTAGATA | ATCAATAACT | TGTTCTAATA | TATAAATTCT | AGCAACTTTA | 480 |
| AACGACATAT | TGCCAATTAC | TGTTTTAGGT | GCAGGTTTCG | TTAATAATGG | CAATAATACT | 540 |
| TGCGCAACTA | CCAAACTAAT | AATAACCATA | CCAGATGCAA | TAAATAATAA | GTCGTTTCTA | 600 |

55

30

35

40

45

| | ATTGTTCCAT | GCACACCACA | TAACGTCATA | ATTAAAGCGT | ATAAACTTCG | CTTTGGTGGT | 720 |
|----|------------|------------|------------|------------|------------|------------|--------|
| | TTCTCAGTCG | TTGGATTATC | ATCATCATTT | TTAGTCATCA | TTTTTTGGAA | TGGACTGATG | 780 |
| 5 | GCTAAATAAA | AATAAGGATA | TAAGACATAA | ACCCAAACAA | ATCTAAATAG | ATAGACAGCT | 840 |
| | AAAGCAACAA | CAATAGTGAT | GCCTATTAAA | AAGATTAAAT | TGTGCGGTTC | TGTTTTGATA | 900 |
| | ATTTTAATAA | TAACTTCAGG | TACTAAAAAT | CCTAATATTG | AAAAAACAAA | GCCATTTAAA | 960 |
| 0 | ACATAACCTA | GTATATTCCA | TGTATGATTG | TAACTCATTT | GCAGTTGTGT | ACGTACTTGC | 1020 |
| | ATAATTCTGT | CACGTTCGAA | ACCATGTACA | AGTCCTGCAA | CTACTGCTGC | AATGATTCCT | 1080 |
| 15 | GATGCGTGaA | ACAATTCAGC | AATTAAATAC | GTAACAAATG | GTGTTAACAA | TTGAATAATT | 1140 |
| .5 | GTAAACATAT | TAATGTTTTC | ATATCCTCGA | CGCATCAATG | TTAATCGGAA | CCTTACTAAT | 1200 |
| | GCCATACCTA | TAAGTAAACC | AACCACTGCG | CCACCAATTG | ATGCAATTAA | AAACAACTGA | 1260 |
| 20 | ACAGCATCAA | CAAGTGAAAA | AGCACCTGTA | ACTAATACTC | CAACAGCTAT | TTTAAATGAA | 1320 |
| | ATAATACCAG | CAGCATCATT | CAATAATGAC | TCACCTTCAA | GAATTGTCAT | TGCTCCTTTT | 1380 |
| | GGCAAGACCT | TTCCTTTAGT | GATTGCTTGC | ACTGCTACTG | CATCAGTAGG | ACAAAGAATG | 1440 |
| 25 | GCAGCAATTG | CAAATGCTGC | TCCAATAGGT | AAATCTGGCC | AAATCCAATG | AATAAATAAA | , 1500 |
| | CCTACACCTA | TCACAGTAGT | AATGACTAAT | CCTAATGCCA | TCATCATCAC | TGGCTTAATA | 1560 |
| | TATTTCCTTA | AATGGACTCT | AGAAACATTA | ACACCTTCTA | CAAATAACAA | AGGCGCAATC | 1620 |
| 30 | ATTGTTACCA | TAAACAATTC | AGAATCAAAA | TTAAATTGAA | CAGGGATTGG | GGTAATAAAT | 1680 |
| | AGTAACATGC | CCAAGAAAAT | TTGTATAAAT | GCTAGGGGTA | CTTTAGGTAT | GAAAGTATGG | 1740 |
| | ACAAACGAAC | TTAGTATCAC | AACAGCTATA | AATATAAGAA | TTGTTTCAAA | TATTTCCAAA | 1800 |
| 35 | CTTTCACCTC | TCTAAAAAGT | ATTGTTTAAT | TGAAAATTAA | GTATCACATC | TCGTTGTAAT | 1860 |
| | TATACTTTAG | AGGATAAATT | GAGTTAGCGA | CCACAAAAGC | ACTTTAATAT | AGATATATGT | 1920 |
| 40 | CTACGATTGC | AGTACTTAAA | TTTGCAATTA | TTTAATTTTA | TTTTATCACT | AATTGTTTGT | 1980 |
| | ATAAATAAAC | AACTTGCTTT | CACATAACAA | CATTAACTTA | TAATACAAAA | AATGAGCACC | 2040 |
| | TTAAAATCGA | CTAACCAATT | TCAAAGTACT | CTTTTAATGA | TTAATTTTGA | AAACAGATTT | 2100 |
| 45 | TCAAAAGCAT | TGTTATGCTT | AACAATTTAG | CCAACACTTC | AATCGTTTTG | ATACCATTTC | 2160 |
| | TTACGATGCT | CTTCTCGTTT | TTCAGCACGT | AATTGTAATG | CTTCTGTAGA | GTTTTGTTCA | 2220 |
| | TTTGAACTTA | ATAATATTGA | TGCATGTGTG | TGAGCATCAT | TTTTTCGATA | CATATAAGCG | 2280 |
| 50 | CCGTTGCGAT | AAGCAGCGCG | AGCGACTAAG | TGCATGCCGA | CTGGTGAAGT | TAAATTAATA | 2340 |
| | AAAACAAGTG | ACACTAATAA | ACCCACACTO | AAAAATCCTG | TATTCACAAT | TAAATAAAT | 2400 |

| | CITAAGAAAA | CATCITGGAA | TTTCACGATA | CCTATIGCAC | TAATAAGAGC | AATAAAACTA | 2520 |
|----|------------|------------|------------|------------|------------|------------|------|
| _ | CCTAACAACA | ACATCACAGC | AGCAATAAGA | CTAAAGATTT | CTTTTGTTAT | TTCCATTAAA | 2580 |
| 5 | CACATGCCCC | CCACCAATAA | AGCGTGATAT | TGAAACAGAA | CTTACAAAAG | ATATAATGGC | 2640 |
| | AATGAGCATG | ATTGAATCTA | AGAAAGAAAC | GGTGCCCATA | AGTACACTTA | ACACACCCAC | 2700 |
| 10 | AATTGACATT | ACGACAGCAC | TTGTTGTATC | AAATGTAACG | ACACGATCTG | CTGTTGTAGG | 2760 |
| | TCCCTTGATT | AATCTAAATA | AACAGATGAT | TAATGCAATT | ССААААТАА | TGAGTGAACT | 2820 |
| | AATAATCATA | ATATGTGTTA | TTGTTTGTAT | CATCGCGACA | CCTCCAATAT | TAAGTCTTCA | 2880 |
| 15 | TAATGCTTAA | TACTTCTTAA | CAAACTATCT | TTTTCTTTTT | CTGACACGTC | GATACTATGA | 2940 |
| | ATAAAAAACT | TTTTAGAGTC | TTGAGAAATT | CGTATTACTG | TAGACCCTGG | AGTTATAATA | 3000 |
| | ATTAAAATTG | TTAAAAATGT | TATTGACCAA | TCACTTGTTA | GTCTTGTTTC | ATATGAAAGT | 3060 |
| 20 | AATCCAGGGT | TCATATCTTT | TGTTTTAAAA | AGAATATAAT | TAATCGTGCT | AATGCTAGAT | 3120 |
| | GTTATTAATT | GATATAAATA | AACACCTAAA | AATTTAATAG | CTACCCATAT | TTTTCTAACA | 3180 |
| | TAAAAATCAT | CGCTGAAAAA | CCTGTGTAAT | ATATAAATGA | CAATTAAACC | AATTAGATAT | 3240 |
| 25 | CCAGAAAAGA | AAGTCGAGAA | TTTAAAATGA | TCTTCATCTT | GAAATAATAC | CCATAAGAAT | 3300 |
| | GCAATGATAA | TATTTAAAAC | TATTTGATTC | ATTTAGTCCT | CTCCTTTCAA | ATGCGGATTT | 3360 |
| | ACAAGTTTTT | GATATAATTG | ATCACTCGTG | TTCAACTCAG | TTGCATCACT | TGTAACATTT | 3420 |
| 30 | AACACAACAG | GTGCAGCAAT | TCCGATTGCG | ATAACCACAA | стастаааат | ACTTAAAATT | 3480 |
| | CTTTTTCGAT | ATAGCGGGAT | TTTCTTAAAA | TTAACTTCCT | CCCCATCTTT | ATCTCCAAAA | 3540 |
| 35 | TACATATAAA | AAAGTATCCT | AAATAAACTG | TACATTGCAA | TTAGACTAGT | AATAATCATT | 3600 |
| | AACGCTAGTC | CAATATAATT | GCCATTTTGC | AATGCACCTT | GGAAAATAAG | TACTTTCCCC | 3660 |
| | GGAĀĀGCCAC | TAAATGGAGG | CACGCCGCCA | ATAGCAAAAA | TCATTATAAT | AAACGCAACT | 3720 |
| 10 | CCAAATAAAG | GTTCTTTTTT | AGCTAAGCCA | TTCAAATATT | GATATTGTCG | ATAGCCTGTA | 3780 |
| | ATGTAAACTA | AACTACCAAT | AATAAAAAAT | AGCAATGTTT | TTACAACAAT | GTCATTTACC | 3840 |
| | AAATAAATA | TTGCACCATT | AATACCTGCA | AACGTGTTTG | TTCCTAAACC | TAAAATGATA | 3900 |
| 15 | AATCCTATTG | AGATTATGAC | TTGGTAAGCT | GCAATCTTTT | TAATATCTTT | ATAAGCAATG | 3960 |
| | ACACCTATAG | CGCCGATGAC | CATAGTTATA | GCAGCCATAG | TTGCTAGCAA | TGGATGTATG | 4020 |
| | AGATCATTAT | GTTGATCAAA | TAGTAAAGTG | AAGAATCGAA | TTAATGCATA | GGCCCCTACT | 408 |
| 50 | TTGGTCATTA | ACGCTGCAAA | TAATGCTGCA | AGCTCAGTAT | TTAACACAGC | GTAGGCTTTG | 414 |
| | CCTACCCACA | TAAAAACCAC | CACCCCTCCT | TTCCCACTAA | ATGCGACTAA | CARCATTART | 420 |

| | AAGTTTAATG | TACCTACTGT | TTTATAAAGT | AAACCTATAC | CTAATAAGAA | TAGCCATGAA | 4320 |
|---------|------------|------------|------------|------------|------------|--------------|------|
| | CCAATAATAT | TCAAGACAAC | ATAAATAATT | GCAGCACGTA | ATTGTTCTAC | AGATTGTCCA | 4380 |
| 5 | AGTGTAATGA | GTACAAATGA | CGCTAGTAAC | ATAATTTCAA | ACATGACGTA | TAAATTAAAT | 4440 |
| | AAATCTGATG | TTAGAAAAGA | GCCTATCACG | CCAACACTTA | AAAATAATAT | GAACGATGGC | 4500 |
| 10 | AAGTGATAAC | GATTTGCTTT | ATGTTCGCCA | CGCCCAAATC | CGTATGCCAT | AATTAAAGTA | 4560 |
| 70 | ATCACAAACG | AAGCGGTTGT | AACCATAATT | AAACTTAAAG | AATCTCCTAA | AAACTGTATA | 4620 |
| | CCAAAGGGCG | CTGACCATCC | TCCAAAGTCT | AGCGTAATTG | GACGGTGACG | CTGAACATAA | 4680 |
| 15 | ATTAATAGCA | TTAATGAAAT | AATTGTGGTG | ATAGTCATTG | TACCTAAGTA | TAAATATTTA | 4740 |
| | GAAATACGAT | CATTATTTTT | TAAAAATACA | AGGATTAAGG | CACAAAGGAA | TGGTAATAAC | 4800 |
| | ATTGGTAAAA | TCAATAAGTT | ACTTAGCATC | ATCTTCCCCC | CTTAGGCCTT | CAATTTCATC | 4860 |
| 20 | TTCTTTTGTT | ACTITATAAG | TTCTATAAAC | AAGTACAAGT | AAAAACGCAG | TCATCCCAAA | 4920 |
| | CCCTATAACT | ATTGCAGTTA | GTACAATAGC | TTGTAACAAG | GGATCAACAA | ACAATTGGTT | 4980 |
| | TCCACCAGTT | ATTAGTGGTT | CTGATCTACT | AGAACCATAC | GTTCCCATAC | TCATAATAAT | 5040 |
| 25 | GAGATTACCA | GCATGAGTAT | ATATTGAAAT | TCCGATTACA | ATACGAATTA | AATTGATTGA | 5100 |
| | TAAAATCATA | TATGTTCCTA | TAAACACTAA | AAATCCTATA | ACTAGTAATA | ATATTAAATT | 5160 |
| | CATGATCGAC | CTCCGCTAAG | CGACAACATC | ACTGTGACAA | TAACACCAAC | AACTGAGAAT | 5220 |
| 30 | AAAATACCTA | ATTCAAAAAG | TGTTATTGTA | CTTACATGAA | TTTGTCCTAA | AATTGGAAGT . | 5280 |
| | ATCCAAGTTG | TTTCATATTG | AGACAAAAAT | GGTTTTCCAA | AAAACATAGG | TATTATCGCA | 5340 |
| 35 | GTAATAGATG | ATACCAATGC | TCCAATAATC | ATTAAAATTC | TAAAATCAAT | CGGTAAACTT | 5400 |
| | TCTAAAACCT | CTTCAACATT | AAAAGCCAGA | AACATTAAAA | TAAACGCTGA | ACTAAATATT | 5460 |
| | AAAÇCACCAA | TAAACCCACC | ACCAGGATTA | TTATGACCTG | CGAAGAAGAC | ATAGAATCCG | 5520 |
| 40 | AAAGTCAATA | AAATAAATAC | AACAAGTTTC | GTGACCGTTC | TTAACACGAC | ATCATTCTCT | 5580 |
| | TTCATCTTGT | CCCCTCCGAT | CTTGATAATT | TAATAATGtg | TAAATACCTA | GCCCAGTAAT | 5640 |
| | AATTAACACT | AATCCTTCAA | ATAATGTATC | TAATGCTCTA | AAGTCACCAA | GTATCGCATT | 5700 |
| 45 | TACAATATTT | TTACCACCTG | TTAGTTTGTC | AGCTTTTAAA | TAAAAGTCTG | ATATTGATGA | 5760 |
| | TAAACCATCT | GTTTGTTGTG | TAATAAAAAT | TAATGATACA | ACAATAAGTG | CCATCAAGAG | 5820 |
| | TGATACAGAA | ATTITAATTA | TITCTCTTTT | TTTGTTAGCG | TTAGATCTTG | GCACGTTTGG | 5880 |
| 50 | TAATCTTGAA | AAACTGACAA | TAAATAGTAT | CGTCGTTATT | GTTTCAACTA | CTAGCTGAGT | 5940 |
| | CAATGCTAGA | TCAGGGGGTT | TCATTGCTAT | AAAGAATAAG | GTCACAACAA | ATCCGATGAC | 6000 |

| | GACAGTTACG | ATTGCTAATA | TAATTTCTAA | TGCCCCAAAT | TCAGAAACAT | GTAACTGATG | 6120 |
|----|------------|-------------------|--------------|--------------|------------|-------------|------|
| | TACTTTAGGA | AGTCCAATTC | GAATATAACC | ATATCCAATG | ATAATCATAA | ATATGCCTAA | 6180 |
| 5 | GGTCATAATA | ATGTACTGGT | TTAAACGATC | TTGCATAACA | CGTTTAAATC | GCTTCGTAGC | 6240 |
| | AAACTTTTCA | AAATGTCGAT | ATACCATCTC | ATAGCTTTTT | GAAACTGAAA | TCTGTCTAAT | 6300 |
| | TTTACCTGTG | AACACTTTTT | TCCAATCTAC | TTTGATTGCT | AGTACACTAC | CCAATAAAAT | 6360 |
| 10 | AATGATGATG | GTTAAAAGAA | GCGGTATGTT | AAATCCATGC | CATTGCGAAA | CATGTGGTGC | 6420 |
| | CAATTGATCA | ATTTGATGAT | TACCACCTGA | TACAGCTCTT | AATGCnAGAA | CGATAATCCC | 6480 |
| 15 | CTTCCCAAAT | ATATNTGGTA | CAAAAAAGAT | TACAGGTACT | AGCACCATTA | aTATAAGAGA | 6540 |
| | TGGTAAACTA | aACAACCATG | GTTCGTGGAT | ATTTTTTTTA | GTAAAAACCT | TAGAATCATA | 6600 |
| | TTTTGtCCAA | AATACTTCTT | TTACCATGTA | TAGTGCATAT | GTGAATGTAA | AAACACTCGC | 6660 |
| 20 | AATAACACCA | ACAAACACGA | TAGCTATCAT | TGAAATCAAA | CTAAATTGGG | ATAATTGTCC | 6720 |
| | AGTTTGTGTT | AATGCATCTA | AAAACATTTC | TTTACTTAAA | AATCCATTTA | AAAATGGTAC | 6780 |
| | TCCAGCCATA | GATAGAGCCG | CTATCGTCAT | GACTAGATTC | ATTTTAGGAA | ATAGTTGACG | 6840 |
| 25 | CATTCCACTT | AAAATTCGTA | TATCCCTTGA | ACCTGCTTCA | TGATCTAAAA | TACCTACTCC | 6900 |
| | CATGAAAAGC | GCACATTTAA | AGATGGCATG | ATTCATTAGa | TGAAATAGcG | CACCATATAA | 6960 |
| | TACMAATACA | TAAATaGATG | CTATTGCGTC | TTGTTGGTGT | TGAGCATATC | CGCCACCTAT | 7020 |
| 30 | ACCCACCATA | GCCATAATCA | TCCCAAGTTG | ACTGATTGTA | GAGTACGCTA | GGATACCTTT | 7080 |
| | TAAATCCCAT | TGTTTTAAAG | CTGTAATTGA | ACCAAATAAC | ATTGTTATTA | AACCAACAAA | 7140 |
| | CGTAACGATA | TATACGTACA | TATTGCTANG | ACCTAATAAT | GGTGTAAATC | GAAGTAATAG | 7200 |
| 35 | AAnGATACCA | GCTTTTACCA | TCGTGGCTGA | ATGTAAATAA | GCACTTACAG | GTGTAGGTGC | 7260 |
| | AGCCÁTTGCT | CTAGGTAGCC | AGTATGAAAT | GGAraTTGTG | CTGATTTTGT | AAATGCACCT | 7320 |
| 10 | AATAAAAACA | TAAAAATCAT | AGGGATAAAC | AATCCATGAT | TCTTAATATG | ATCTGCTTGT | 7380 |
| | CCTAATATCT | CTGTGATGTT | ATTCGTTCCT | GTCATGATAT | ACAGCATAAT | AAAACCAACT | 7440 |
| | AATAACGCCA | ATCCACCAAA | TACTGTAATC | ATAAATGATT | GAATCGCACC | AAATTGACTG | 7500 |
| 15 | TCACCATTGT | TATACCAATA | TGANATCAAT | AAAAATGATG | ATmCACTCGT | TAATTCCCAA | 7560 |
| | AAaATGTACA | TCmATATCGT | ATTGTCTGAT | AATACaaTAC | CAATCATACT | GAACATAAAT | 7620 |
| | AACGTTAAAT | AAAATAAAA | CCTTGGTAAA | TTGTCTTTTC | GAGAGGATAA | ATATTGAGTT | 7680 |
| 50 | GCATAGAAGA | ATACTGCAAT | TCCAATAAGT | GAAATAATAA | GAGAAAACAT | TAAACTTAAA | 7740 |
| | CCATCTAAAC | A A TOTA A A A TO | מייים מיייים | A ATCTCTTA A | TCCNTCCNNT | 1C1CCC11C11 | 2000 |

| | | _ |
|-----------|--|------|
| | GGTGCAACCA ACGCTATGTA CCCGGCATAT TTAGCCAATG CTCTACGTTT AGACATTAGA | 7920 |
| | AGTATCATCG CCATAATCAC AAGTATAGCA ATTAATAAAT AAACCAAACT CATTATTAGC | 7980 |
| 5 | CTCCTTTGTT TCTATAATTG TAATGAAATA TAAATACTAT GTTCACACTC ATTTTCTAAA | 8040 |
| | CCGATAAAAT TTAGTGTTTC AATAGCAGAT TGATGCCCTA AATACTTTTG AATGACTGGT | 8100 |
| 10 | ATAAGTATAC CTTTTTGATA AGCATGATAT GCAAATGTCT TACGCAATGT CGTTAGTCCT | 8160 |
| ,,, | ACATTATCTA TACCAGCTTC AATTGATGCT TGGTGAATTA TTCGATATGC TTGCTGTCTA | 8220 |
| | GATAATACTT GATTTGTTCG TAGTGATTGA AAAAGAACGT CTTCATTCGA AAGACTCCTG | 8280 |
| 15 | TCCTCTATAT ATTGAAGTAG TTCTTTCGAT AATGTTTCTG GTAACCTAAT TTTAATCAA | 8339 |
| | (2) INFORMATION FOR SEQ ID NO: 176: | |
| 20 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 588 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 25 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 176: | |
| | CCCGATTTT TTACGTAATC TAATACATAC GGCAAAATCA ACTTTAATCA AAAAAGACTC | 60 |
| 30 | ATACACAATG CCTTTAAAGC ACATGTATGA GTCCTTTTTA GTAGTTTATA TCAAAAAATA | 120 |
| 30 | GTTTAATGTA TAAATTAGTT TTTGTTTACA GATGCGTCGT AGATTGATTC TACAGCATCA | 180 |
| | CCTAAAGCTT TATCGAATTC TTCTTTAGAT TGATCAGCTC TTAAATCACT AGCTAATGCA | 240 |
| 35 | CGTGAGAAAC TTGCGATAAG TTCAGCGTTA TCTTTAAGTA ATTCATTTGC TTTTTCTCTG | 300 |
| | CTGTAACCAC CTGATAATAC AACGACACGA ACAACATTAG GATGTTCAGC TAACTCTTTG | 360 |
| | TATAAGTTTG GTTCAGTAGG AATTGTTAAT TTCAACATTA CTAATTGATC AGCATTTAAG | 420 |
| 40 | CTATCTAAAC CTTTTTTAAG TTCAGCTTTT AATACTTTTT CAATTTCAGC TTTGTCTTTT | 480 |
| | GCATTAATAT TAACTTCTGG TTCGATAATT GGAACTAAAC CTTTAGCAAT AATTTGTTTA | 540 |
| | GCAACTTCAA ATTGTTGTTC AACAACGTCT TTGATACCTT GCTCATTT | 588 |
| 45 | (2) INFORMATION FOR SEQ ID NO: 177: | |
| 50 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2841 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |

| | ATAGAGTNCT GGNACTTACT ATGACATATG GCGCTAGAAT GGCTGAGCCA GGTGAATTTA | - 60 |
|----|---|------|
| | CAAAACGTGC CTTTTTAAAT GGTCGTATTG ATTTATCTCA AGCTGAAGCA GTTATGGACT | 120 |
| 5 | TTATTCGCTC GAAGACAGAT AGAGCTTCTA AAGTTGCGAT GAATCAAATT GAAGGTCGTC | 180 |
| | TAAGTGACTT AATCAAAAAA CAACGTCAAT CTATATTAGA GATACTCGCT CAAGTGGAAG | 240 |
| | TGAATATTGA TTATCCTGAA TACGATGATG TTGAAGATGC GACTACTGAA TTTCTTTTAG | 300 |
| 10 | AGCAGTCTAA AGAAATCAAA CAGGAAATTA ATCGTTTATT AGATACCGGT GCGCAGGGTA | 360 |
| | AAATTATGCG TGAAGGTTTA TCTACAGTTA TTGTTGGTAA ACCAAACGTA GGTAAATCAT | 420 |
| | CGATGTTAAA TAATTTAATA CAAGATAATA AAGCGATTGT AACTGAGGTA GCAGGTACTA | 480 |
| 15 | CTAGAGATGT CTTAGAAGAG TACGTCAATG TTCGTGGCGT GCCATTAAGA TTAGTTGATA | 540 |
| | CTGCTGGTAT ACGTGAGACA GAAGATATAG TAGAGAAGAT TGGTGTTGAA CGCTCTAGAA | 600 |
| 20 | AGGCTCTTAG CCAAGCAGAC TTAATTTTAT TTGTATTAAA CAATAACGAA GCATTGACWC | 660 |
| 20 | AAGAAGATTA CACATTATAT GAAGTGGTTA AAAATGAAGA TGTAATCGTA ATTGTTAATA | 720 |
| | AAATGGATTT AGAGCAAAAC ATAGATATTA ATGAAGTTAA AGATATGATA GGTGATACGC | 780 |
| 25 | CATTAATTCA AACTTCAATG TTAAAACAAG AAGGTATTGA TGAATTAGAA ATACAAATTC | 840 |
| | GAGATTTGTT CTTTGGTGGA GAAGTACAAA ATCAAGATAT GACTTATGTT TCTAATTCAA | 900 |
| | GACATATTTC ATTATTAAAA CAAGCAAGAC AAACGATACA AGATGCGATT GATGCAGCAG | 960 |
| 30 | AATCTGGTGT GCCTATGGAT ATGGTACAAA TTGATTTAAC TAGAACTTGG GAAATATTAG | 1020 |
| | GAGAAATTAT TGGTGAGACT GCAAGTGATG AACTCATCGA TCAGTTATTC AGTCAATTCT | 1080 |
| | GCTTAGGTAA ATAGTAATTG AAATAGACGG AATACCGTCT TAAGAAGGCT AGTAAGATAT | 1140 |
| 35 | CAAATAAGGA GGTTTATATT GTGGTTCAAG AATATGATGT AATCGTTATA GGTGCGGGAC | 1200 |
| | ATGCAGGTGT AGAAGCAGGT TTAGCATCTG CAAGACGTGG TGCTAAAACA TTAATGCTAA | 1260 |
| | CAATAAATTT AGATAATATT GCATTTATGC CATGTAACCC ATCTGTAGGT GGACCAGCTA | 1320 |
| 10 | AAGGTATCGT TGTTCGTGAA ATTGATGCTT TAGGTGGACA AATGGCAAAA ACAATCGATA | 1380 |
| | AAACACACAT TCAAATGAGA ATGTTAAATA CAGGTAAAGG ACCTGCTGTA AGAGCACTAA | 1440 |
| | GAGCGCAAGC AGATAAAGTA CTTTATCAAC AAGAAATGAA ACGCGTGATT GAAGATGAAG | 1500 |
| 5 | AAAATTTGCA TATAATGCAA GGTATGGTAG ACGAACTTAT TATAGAAGAT AATGAAGTTA | 1560 |
| | AAGGTGTACG TACAAATATT GGTACAGAGT ATTTATCTAA AGCAGTAATT ATTACAACGG | 1620 |
| | GAACATTTTT ACGTGGTGAA ATCATTTTAG GTAATATGAA GTATTCAAGT GGACCAAATC | 1680 |
| 0 | ACCAATTACC ATCAATCACA TTATCAGACA ATTTAAGAGA ACTTGGTTTT GATATTCTTC | 1740 |

| | AAATACAACC | AGGTGACGAT | GTAGGTCGTG | CATTCAGCTT | TGAAACAACA | GAATATATAT | 1860 |
|----|-------------|--------------|-------------|------------|------------|------------|------|
| | TAGATCAATT | GCCATGTTGG | CTAACGTATA | CTAATGCTGA | AACACACAAA | GTTATCGATG | 1920 |
| 5 | ATAATTTACA | TCTATCTGCA | ATGTATTCAG | GGATGATTAA | AGGAACCGGG | CCACGTTATT | 1980 |
| | GCCCTTCAAT | TGAAGATAAA | TTTGTTCGAT | TTAATGATAA | GCCGCGACAT | CAACTTTTCT | 2040 |
| | TAGAGCCTGA | AGGTCGTAAT | ACAAATGAAG | TATATGTGCA | AGGATTGTCT | ACAAGTCTTC | 2100 |
| 10 | CTGAACATGT | GCACGTCAAA | TGTTAGAGAC | GATACCAGGT | CTTGAAAAAG | CAGATATGAT | 2160 |
| | GCGTGCCGGC | TACGCAATTG | AATATGATGC | GATTGTGCCA | ACGCAGTTAT | GGCCTACACT | 2220 |
| | TGAAACGAAA | ATGATTAAAA | ACTTATATAC | TGCAGGTCAA | ATTAATGGTA | CATCTGGTTA | 2280 |
| 15 | TGAAGAAGCA | GCAGGACAAG | GATTGATGGC | AGGTATTAAC | GCTGCAGGTA | AAGTGTTAAA | 2340 |
| | CACAGGCGAA | AAGATATTAA | GTCGTTCAGA | TGCATATATT | GGTGTCTTAA | TCGATGATCT | 2400 |
| | TGTAACTAAA | GGTACTAATG | AACCTTATCG | TTTACTAACA | TCACGTGCAG | AATATCGTTT | 2460 |
| 20 | GTTACTACGT | CATGATAATG | CTGATTTGAG | ATTGACGGAT | ATGGGATATG | AACTTGGTAT | 2520 |
| | GATTTCTGAA | GAAAGATATG | CACGTTTTAA | TGAAAAACGT | CAGCAAATTG | ATGCGGAAAT | 2580 |
| 25 | TAAGCGTTTA | TCAGATATTC | GTATTAAACC | AAACGAACAT | ACGCAAGCGA | TTATTGAACA | 2640 |
| | ACATGGTGGT | TCTCGCTTAA | AAGATGGTAT | TTTAGCTATC | GATTTATTAC | GCAGACCTGA | 2700 |
| | AATGACTTAC | GATATAATTT | TAGAACTTTT | AGAAGAAGAA | CATCAATTGA | ATGCAGATGT | 2760 |
| 30 | TGAAGAACAA | GTAGAAATAC | AAACAAAATA | TGAAGGTTAT | ATCAATAAAT | CACTACAACA | 2820 |
| | AGTTGAGAAA | GTTAAGCGTA | т | | · | | 2841 |
| | (2) INFORMA | ATION FOR SE | Q ID NO: 17 | 78: | | | |
| 35 | (i) SE | EQUENCE CHAR | ACTERISTICS | S: | | | |

- (A) LENGTH: 3025 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

- (D) TOPOLOGY: linear

40

45

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 178:

| ATCTAATTTC | AAACCCGGTG | ATAAATTGCC | AAGCGTGACG | CAATTAAAAG | AACGTTATCA | 60 |
|------------|------------|------------|------------|------------|------------|-----|
| agtaagtaag | AGTACTATCA | TTAAAGCATT | AGGCTTATTG | GAACAAGATG | GTTTGATCTA | 120 |
| TCAAGCACAA | GGCAGTGGTA | TTTATGTGAG | AAATATTGCT | GATGCCAATC | GTATCAACGT | 180 |
| CTTTAAGACT | AATGGTTTCT | CTAAAAGTTT | AGGTGAACAC | CGAATGACAA | GTAAGGTACT | 240 |
| TGTTTTTAAG | GAGATTGCAA | CGCCACCTAA | ATCTGTACAA | GATGAGCTCC | AATTAAATGC | 300 |

| | CGAATATTCT TATTATCATA AAGAAATCGT GAAATATTTA AATGATGATA TTGCTAAGGG | 420 |
|----|--|------|
| 5 | CTCTATCTTC GACTATTTAG AATCAAACAT GAAACTTCGT ATTGGTTTTT CAGATATTTT | 480 |
| | CTTTAATGTA GATCAACTCA CTTCAAGTGA AGCTTCATTA CTACAATTGT CTACAGGTGA | 540 |
| | ACCATGTTTA CGTTACCACC AGACTTTTTA TACAATGACT GGCAAACCCT TTGATTCATC | 600 |
| 10 | TGACATCGTA TTTCATTATC GTCATGCACA GTTTTATATT CCTAGTAAAA AGTAATAAAT | 660 |
| | ACATAAAAAC GTCTATATCC CAGTTATAAA CTGGAGTATA GACGTTTTTT TACGATAATA | 720 |
| | ACAATGGCTC AAATTGCTAT TATCTTGCTT AGGTTTTTCG TTTTAGAAGA ATATTGCTAC | 780 |
| 15 | AAAGACAGGC ACAACTGCTA CAACAACTAC ACCAACTAAC ACTAAAGCTA TACTTGCCAT | 840 |
| | TGATTCTTCT ACAGGTCCTA ATTCTTTGGC TGGTGCTACA CCTAATGTGT GACCACTTGT | 900 |
| | TCCAAGTGCT AATCCTCGGG CAATAGGGTT AGTAATTCGG AAAAGCTTTA AGAATTTATT | 960 |
| 20 | ACCTAGGGCA TAAATAATGA CACCATTTAA AATAACTGCT AATGATGTTA ATTCTTTTAT | 1020 |
| | ACCACCGATA CCAGCTGATA CTGGTAACGC AATCGCTGTA GTTGCTGCTT GAGGTAACAT | 1080 |
| | TGATAAAATA ACATCATTGG CAAATTGTGC TAACTTCGCA AAAGTTAAAA TAATTAATAA | 1140 |
| 25 | CGCTACAACT GTACCGATAC CAATACCTCC GATGATACGA TGCCAATGTT TAACAAGCAC | 1200 |
| | TTCACGCTTT TTATATAACG GAATCGCAAA ACAGATTGTT GCCGGTTCTA AGAAGAAGTA | 1260 |
| | AATAATGTCT CCACCTATTI TGTAAGTCTT ATACGGAATG CCTGTTAAAT AGAGGAAGGC | 1320 |
| 30 | CACACCAAAT ACCATACTGA CAAATAGCGG TGCGAATAAG AAGAAACGAT TAGTTTTTTC | 1380 |
| | AAATAATATG GTCGCTAAGA AAAATGGTAT AACGGATAAC AGTATTCCGA AGTAAGGTGT | 1440 |
| | GTTTAGTGCT AAGTGGTTAA TCATGAGCTT GTGCCTCCTC TATTTTGATC TTTTTTGTGA | 1500 |
| 35 | CTTTGTCACC TTTAGATCTC GAAGTAACTT TCATAATAAT TTGTGTGACA TAGCCAGTAC | 1560 |
| | AAATÄAGTAA TAGTATTGTT GAGACGATTA TTAGTCCAAT GATTAAAAAT GGTGCTTGGC | 1620 |
| 10 | TAATGACACC TAAAGAGTTA ACAACTGAGA TACCGGCTGG TACGAAGAGT AAGCCAATGT | 1680 |
| | TATITGTTAG TGTCGTTCCT ACTITTCGA CTTCGCCTAA CTTAACAGCA CCAGTACATA | 1740 |
| :5 | ATAATACAAA TAATAATACT AAACCGATTA CTGATGCAGG CATAGGAATT GGCATAAATG | 1800 |
| | ATTCAATTAT TTTCGATACA AAGAGTACTA AAGCAATTAC AATGACTTGG TGAAAAAAGT | 1860 |
| | GTGCTGGTTT TGATGCGTCT TTTTGTTGTT TCACGACCAT TGCCTCCTAC GTTTGATTTA | 1920 |
| | ACTAAAGTAT AGATGGCTCA CTTCGATTTG CGTGATTTTT AGTCCGAAAT ACAAAATATC | 1980 |
| ю | ATAGGTAAAA TGCATAAAAA AAAGGATTAC TGTTAAAAGTA ATCCTATCGA CGCTTTAAAA | 2040 |
| | TCTTTCATAA ATGAACGTCC AACTTGCATC TTGACACCAT TTGTCAATAT TAGGATAT | |

| | TGAATACGTA | TAAAATAAGT | GGGATTCAAT | CGTTTTTCAT | AACGATTCAA | TGGCTCTGTT | 2220 |
|----|------------|------------|------------|------------|------------|------------|------|
| | GTTTCGTATT | TATGATTCGT | TGTATGTATG | GTTGTAATAC | CATTATGTGT | GCCAATCCCA | 2280 |
| 5 | TTTTATAATA | GTTGCTTTAA | CATGTGAATT | TTATCGTCAA | TTTCAACAGG | TAAGCTTTGA | 2340 |
| | TCAAAATTCG | CCGACATATC | ATTCGCAATT | GCACTTGCGT | TATTATCATC | TTTGGCTTTA | 2400 |
| | GTCGCACGCA | CTTTATTGAC | TGCTTGTTCA | ATACGTTTTT | GACCAAACGG | TTTCAAAATA | 2460 |
| 10 | TAGTCTGTCG | CATTTAATTC | AAATGCCTGT | ACTGCGTATT | GGTCATGTGC | AGTTGCAAAA | 2520 |
| | ATAATCGCAG | GTGGCTCTTT | CATCTTTTGA | ATCTTAGCTC | CTAATTCGAT | CCCATTTTCA | 2580 |
| 15 | TCCATTAAAT | TGACATCTAA | AAATATAATG | TCATATTGAT | TGATCAGTAG | TGCTTCCAAT | 2640 |
| 15 | GTTTCTTTTA | CATTTTCTGC | CTCATTAATT | TCTTCAAAAC | CACCAATTTC | ATTTAATAAA | 2700 |
| | TATGTTAATT | CATTACGTGC | TAATGGCTCA | TCATCTATGA | TTAATGCTTT | CATATTTATT | 2760 |
| 20 | CCTCCTCTTG | TCTTTCATAA | GGAAGTACAC | ACCAAAAAGT | GGTACCGCTC | GATGTCGATT | 2820 |
| | CAAATTGTAA | TGCTGCGGAT | TTTCCAAATA | ATCCTTTTAG | GCGTAAGTTT | AAATTTTCTA | 2880 |
| | AAGCACTACC | AGTTCCAGAC | TCTGATTCTA | CAGATGTnTC | TCCCAACAAA | TGCATTTTAT | 2940 |
| 25 | CTTTAGAAAT | ACCCTGACCA | TTATCTTGTA | CAATAATACG | TACATGTGTT | GCAGTTTCTT | 3000 |
| | TAATCACTGA | CACGTCAATA | TCGTT | | | | 3025 |

(2) INFORMATION FOR SEQ ID NO: 179:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1689 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

35

40

45

50

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 179:

| ACAGAATTTC | ACAGCATTTT | TAGATGAAAA | AATAAGCCAG | TCATAGCGTT | GATTTAACAA | 60 |
|------------|------------|------------|------------|------------|------------|-----|
| ATGAATATCA | AAATTTAGTG | GCTTTATATC | AATAAAGGGT | TTGTGAATAA | TTGATACTAA | 120 |
| ATCACTTTGC | ATGTCATTTG | TTTGTGTCAT | AACTACAACT | GGCTTCATAT | TTAAACGTCA | 18 |
| CICCATTATT | TAATGTTGTT | CATTTAAGCG | TTTTATAATT | TCATAAGCAC | CTTGCTCTTT | 24 |
| TAATTTGTTA | CTCACTGTTT | TGCCTAACTC | AACCGGATCT | GTTCCGTTCA | TTGTATATTC | 30 |
| AAATCGTTCT | TTACCATCTG | GGGTCATAAT | TAAACCTGTA | AATTCGATTT | CGTTTTGATC | 36 |
| TGAGATTGTA | GCATATCCTG | CAATTGGCAC | CTGACAACTA | CCATCCATTT | CTGCTAAAAA | 42 |
| CGTTCGTTCA | GCAGTCACAC | ATTTTGCAAC | CTCATCATTA | TGTACTTTGC | TTAATAATGT | 48 |

| | TAACAATGTA | TCTCTATCAA | GATAAGATGT | TnCAATATCA | TCTGACCAGC | CCATTCTTCT | 600 |
|----------|-------------|-------------|-------------|------------|------------|------------|------|
| _ | TAAACCAGCT | GCAGCTAAAA | TAATCGCATC | ATAATCTTCA | GTTTGTAACT | TTTCTAATCG | 660 |
| <i>5</i> | TGTATCTATA | TTACCTCTAA | TCCATTTAAT | CTCTAAATTA | GGATACTTAG | ATAATATTTG | 720 |
| | TGCACCACGA | CGTAATGAAC | TAGTACCAAT | AATACTGCCT | TCTGGCAATT | GGGATAGTGG | 780 |
| 10 | TGTATGTGTT | TTAGAAATAT | ACGCATCAAA | AGGTAATTCT | CTATCAGGGA | TACAACCTAA | 840 |
| ,,, | TGTTAAACCT | TCCGGAATTA | CACTTGGTAC | GTCTTTAAGC | GAGTGTATTG | CCATATCGAT | 900 |
| * | ATTTTTTCA | AAAAGTTCAT | GTTGTATTTC | TTTÄACAAAT | AAGCCTTTGC | CTCCGACTTT | 960 |
| 15 | AGACAATTGT | TTATCTACTA | TACGATCGCC | TTTCGTGaCA | ATTTCTTTAA | TTTCAATTTC | 1020 |
| | TAGATTTGGC | TCGACAGCTT | TTAATTTATC | AATAAATTGC | TGGCTTTGTG | TTAAAGCTAA | 1080 |
| | TTTACYTCTT | CTGGAGCCAA | CGACTTATTT | ACGCATGTTC | AATTCCTCCT | AGGAACGGAT | 1140 |
| 20 | TGCTCTAGAT | TATTTTCTCA | ATTCACAAAA | TGTGTTGCAA | AAAATAAATT | AATCATATTT | 1200 |
| | AAGCAAAATA | AAATAATGTT | ATAGTATATT | AAATATCTTG | AATTCAACCA | TTTGTTGATT | 1260 |
| | CTAAGTAAAA | TATAACTTCC | АТАТААТАСТ | GTAATAATTG | AAGAGAGTAT | TACCTTCGGG | 1320 |
| ?5 | TCAATGAATA | TACGTTCACC | AACTGAAATT | ACACCCCACT | GTGTACCTAA | AATAATACTA | 1380 |
| | AATATGAGAA | TTATCCACCC | ACTTAACGTT | GAGTAAAACA | CAATTGATTC | AAGTGTAGCA | 1440 |
| | ACGCTACCAA | TTCTAAAGTA | TTTTTGATCA | AAACGTTTTT | CCTTCAAATT | ACGGTATTGC | 1500 |
| 30 | ATGATATACA | GTAATGCATT | GACAAAAGCT | AAGGCAAAGA | AGACATAACT | TAACACAGCT | 1560 |
| | AGACCGATAT | GGACTAACAG | TAACTCGTCT | ACAACAGCAA | TTTTCTGAAC | CTTATTAGTA | 1620 |
| | TAATGTGTCG | GTTGAAATGT | ATTCATCCCT | AAnAGTGTTA | ACCCTATTAA | ATTCCAAGGA | 1680 |
| 35 | AAAACACAG | | | | | | 1689 |
| | (2) INFORMA | TION FOR SE | Q ID NO: 16 | 30: | | | |
| | - | | | | | | |

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1209 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

45

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 180:

NTGGNTGGCT TTTCCTATTG GACCAAATGG ACCNTTTACC TGGCCNTTCC CAGGACACCC 60 CGCTTGTGCC CACATTCCAA TCGGAAAAGG TGTATGTGGT ACAGCCGTTT CAGAACGTCG 120 TACACAAATT GTAGCTGATG TTCATCAATT CGAAGGACAT ATCGCTTGTG ATGCTAATAG 180

55

| | CGATGCCCCT | ATAACGGATC | GATTTGATGA | CAATGACAAa | GAaCATCTTG | AAGCAATTGT | 300 |
|----|------------|------------|------------|------------|------------|------------|------|
| _ | TAAAATTATT | GAAAaGCAAC | TCGCATAAAA | GGACATCAGC | ATTTTCAATA | AAGTGTTGAC | 360 |
| 5 | AGTTAGCAGG | AAAATGTTAC | AATAATCTTT | GTGTGAATTA | ACGAAAGTAG | CAGTTGTATA | 420 |
| | TTATTGAGCG | CTATGTTGTT | CCCAATGCGG | ACGTGTCACG | TAACTGTCGC | TATAAGGTGA | 480 |
| 0 | AGACACATAA | AACAATATAT | CTTAGTAAGC | ATGCAACACT | CTTTTTTGTT | TATTCATAAC | 540 |
| Ü | AACAAAAAAG | AATTAAAGGA | GGAGTCTTAT | TATGGCTCGA | TTCAGAGGTT | CAAACTGGAA | 600 |
| | AAAATCTCGT | CGTTTAGGTA | TCTCTTTAAG | CGGTACTGGT | AAAGAATTAG | AAAAACGTCC | 660 |
| 5 | TTACGCACCA | GGACAACATG | GTCCAAACCA | ACGTAAAAAA | TTATCAGAAT | ATGGTTTACA | 720 |
| | ATTACGTGAA | AAACAAAAAT | TACGTTACTT | ATATGGAATG | ACTGAAAGAC | AATTCCGTAA | 780 |
| | CACATTTGAC | ATCGCTGGTA | AAAAATTCGG | TGTACACGGT | GAAAACTTCA | TGATCTTATT | 840 |
| 20 | AGCAAGTCGT | TTAGACGCTG | TTGTTTATTC | ATTAGGTTTA | GCTCGTACTC | GTCGTCAAGC | 900 |
| | ACGTCAATTA | GTTAACCACG | GTCATATCTT | AGTAGATGGT | AAACGTGTTG | ATATTCCATC | 960 |
| | TTATTCTGTT | AAACCTGGTC | AAACAATTTC | AGTTCGTGAA | AAATCTCAAA | AATTAAACAT | 1020 |
| 25 | CATCGTTGAA | TCAGTTGAAA | TCAACAATTT | CGTACCTGAG | TACTTAAACT | TTGATGCTGA | 1080 |
| | CAGCTTAACT | GGTACTTTCG | TACGTTTACC | AGAACGTAGC | GAATTACCTG | CTGAAATTAA | 1140 |
| | CGAACAATTA | ATCCGTTGAG | TACTACTCAA | GATAATACGG | TCAATACCAA | CACCCACAAT | 1200 |
| | TGTGGGTGT | | | | | | 1209 |

(2) INFORMATION FOR SEQ ID NO: 181:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 698 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

40

45

50

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 181:

AAATCCCTTE GTEAAAGTSC AAAETTTTCC AACTGCTTTA AEAFGACCCA TATTACCETC 60
TTGGATTAAA ECMAGGAATG ACATACCACG ACCACGTATC TTTTAGCAAT ACTTACAACT 120
AAACGTAAGT TCGCTTCTGC AAGTCTTGAT TTTGCTACTT CATCACCTTG TTCAATACGT 180
TTGGCTAATT CGATTTCTTC TTGTGCACTT AATAAGTTAA CACGCCCAAT TTCTTTAAGG 240
TACATACGAA CTGGGTCATT TATTTTAACA CCTGGAGGGG CACTAAGATC ACTTGGATTC 300
AGTTTCTCGT CAGTATCTGA ACTATCTTTT TCATTAACTA GTGAAATATC ATTATCATTT 360

| | | _ |
|----|--|------|
| | GCAATTTCTT CATGACTTAA ATGACCCTCT TTTTTACCTT TTTCAATTAA TTGCTTCTTA | 480 |
| _ | ACATCTTCTA ATGTTAATGT CGGATCAATT GTTTGTTTTT TAATTTTAAC TGTGTTATCA | 540 |
| 5 | GACATGAAAC GGCCTCCCGA TTTTAAATAT GAACATTCGA AATTTATTCA ATATTGCTAT | 600 |
| | TTTAAACGAA ATTCTTAATT AATTCCATCC ATATTTTMAA TTTTATTTTA | 660 |
| 10 | ACTAAATCCC CAATATTTAT TTTTCAATAG TGGTGGTT | 698 |
| 70 | (2) INFORMATION FOR SEQ ID NO: 182: | |
| 15 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 5147 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 20 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 182: | |
| | ACTTGATGAT GTATACAATG TATTTCAAGA ATATTATCAA AAAACATCTA ACATTAAGTT | 60 |
| | TTGTAGAATT CACAATTCTA GCTATTATCA CTTCTCAAAA TAAAAACATC GTTCTTCTTA | 120 |
| 25 | AAGATTTAAT TGAAACAATC CACCATAAAT ACCCTCAAAC TGTTAGAGCT CTCAATAATT | 180 |
| | TAAAAAAGCA AGGCTATCTA ATAAAAGAAC GCTCAACTGA AGATGAAAGA AAAATTTTAA | 240 |
| | TTCATATGGA TGACGCGCAG CAAGACCATG CTGAACAATT ATTAGCTCAA GTGAATCAAT | 300 |
| 30 | TATTAGCAGA TAAAGATCAT TTACATCTTG TTTTTGAATA ATATCTCTAT TACGCAAGTG | 360 |
| | TGCTGTATTC TAAAGTGCAC TTGTGTTTTC TATTTTTTAA TAAAACCTCA GCACATAATG | 420 |
| | AACAACTTTC TATTTTCTAT ATCACTTAAA ACCATTTCCG AAATTAAACC TCAGCACATT | 480 |
| 35 | CAAAGCCCCA CTTTATTCTT AAAAATATTT TTTAACTCAT ATGTATTAAA CCGCTTTCAT | 540 |
| | TATAAAAAT ATCTCTATAT TETATCTGET TETATTAATC GAAATAGCGT GATTTTGCGG | 600 |
| 40 | TTTTAAGCCT TTTACTTCCT GAATAAATCT TTCAGCAAAA TATTTATTTT ATAAGTTGTA | 660 |
| | AAACTTACCT TTAAATTTAA TTATAAATAT AGATTTTAGT ATTGCAATAC ATAATTCGTT | 720 |
| | ATATTATGAT GACTTTACAA ATACATACAG GGGGTATTAA TKTGAAAAAG AAAAACATET | 780 |
| 45 | ATTCAATTCG TAAACTAGGT GTAGGTATEG CATCTGTAAC TTTAGGTACA TTACTTATAT | 840 |
| | CTGGTGGCGT AACACCTGCT GCAAAtgctG CGCAACACGA TGAAGCTCAA CAAAATGCTT | 900 |
| | TTTATCAAGT CTTAAATATG CCTAACTTAA ATGCTGATCA ACGCAATGGT TTTATCCAAA | 960 |
| 50 | GCCTTAAAGA TGATCCAAGC CAAAGTGCTA ACGTTTTAGG TGAAGCTCAA AAACTTAATG | 1020 |

55

1080

ACTCTCAAGC TCCAAAAGCT GATGCGCAAC AAAATAACTT CAACAAAGAT CAACAAAGCG

| | AAAGTCTTAA | AGACGACCCA | AGCCAAAGCA | CTAACGTTTT | AGGTGAAGCT | AATTAAAAAA | 120 |
|-----|------------|------------|------------|------------|------------|-------------|------|
| | ACGAATCTCA | AGCACCGAAA | GCTGATAACA | ATTTCAACAA | AGAACAACAA | AATGCTTTCT | 126 |
| 5 | ATGAAATCTT | GAATATGCCT | AACTTAAACG | AAGAACAACG | CAATGGTTTC | ATCCAAAGCT | 1320 |
| | TAAAAGATGA | CCCAAGCCAA | AGTGCTAACC | TATTGTCAGA | AGCTAAAAAG | TTAAATGAAT | 1380 |
| | CTCAAGCACC | GAAAGCGGAT | AACAAATTCA | ACAAAGAACA | ACAAAATGCT | TTCTATGAAA | 1440 |
| 10 | TCTTACATTT | ACCTAACTTA | AACGAAGAAC | AACGCAATGG | TTTCATCCAA | AGCCTAAAAG | 1500 |
| | ATGACCCAAG | CCAAAGCGCT | AACCTTTTAG | CAGAAGCTAA | AAAGCTAAAT | GATGCTCAAG | 1560 |
| 15 | CACCAAAAGC | TGACAACAAA | TTCAACAAAG | AACAACAAAA | TGCTTTCTAT | GAAATTTTAC | 1620 |
| . • | ATTTACCTAA | CTTAACTGAA | GAACAACGTA | ACGGCTTCAT | CCAAAGCCTT | AAAGACGATC | 1680 |
| | CTTCAGTGAG | CAAAGAAATT | TTAGCAGAAG | CTAAAAAGCT | AAACGATGCT | CAAGCACCAA | 1740 |
| 20 | AAGAGGAAGA | CAATAACAAG | CCTGGCAAAG | AAGACAATAA | CAAGCCTGGC | AAAGAAGACA | 1800 |
| | ACAACAAGCC | TGGTAAAGAA | GACAACAACA | AGCCTGGTAA | AGAAGACAAC | AACAAGCCTG | 1860 |
| | GCAAAGAAGA | CGGCAACAAG | CCTGGTAAAG | AAGACAACAA | AAAACCTGGT | AAAGAAGATG' | 1920 |
| 25 | GCAACAAGCC | TGGTAAAGAA | GACAACAAAA | AACCTGGTAA | AGAAGACGGC | AACAAGCCTG | 1980 |
| | GCAAAGAAGA | TGGCAACAAA | CCTGGTAAAG | AAGATGGTAA | CGGAGTACAT | GTCGTTAAAC | 2040 |
| | CTGGTGATAC | AGTAAATGAC | ATTGCAAAAG | CAAACGGCAC | TACTGCTGAC | AAAATTGCTG | 2100 |
| 30 | CAGATAACAA | ATTAGCTGAT | AAAAACATGA | TCAAACCTGG | TCAAGAACTT | GTTGTTGATA | 2160 |
| | AGAAGCAACC | AGCAAACCAT | GCAGATGCTA | ACAAAGCTCA | AGCATTACCA | GAAACTGGTG | 2220 |
| | AAGAAAATCC | ATTCATCGGT | ACAACTGTAT | TTGGTGGATT | ATCATTAGCC | TTAGGTGCAG | 2280 |
| 35 | CGTTATTAGC | TGGACGTCGT | CGCGAACTAT | AAAAACAAAC | AATACACAAC | GATAGATATC | 2340 |
| | ATTITATCCA | AACCAATTTT | AACTTATATA | CGTTGATTAA | CACATTCTTA | TTTGAAATGA | 2400 |
| | TAAGAATCAT | CTAAATGCAC | GAGCAACATC | TTTTGTTGCT | CAGTGCATTT | TTTATTTTAC | 2460 |
| 40 | TTACTTTTCT | AAACAACTTC | TGAAACGCCT | CAACACTTTC | TACTCTGATT | ACATATATGA | 2520 |
| | CATTTTTAGG | CATTAAAAAA | TCGAACTAGA | CAAGATGCTC | ATTGCATTTC | GTACTAGTTC | 2580 |
| 45 | GATTCATGAA | TAATTAGATT | TAAAATGTCA | TTTGAATCCA | AGTGACAACA | TTATTTATAT | 2640 |
| •• | TTAGAATATT | AACGTTAGTA | TAAACGTCCA | AACACAAATA | AAAGCAACAA | ATATAATACT | 2700 |
| | GTATTTTAAC | GTCATTTTTA | ATAATGCAGA | TTCTTCACCA | ACTTTTTTAA | CAGCTGCAGT | 2760 |
| 50 | CGCAATGGCA | ATTGATTGTG | GTGAAATAAG | TTTCGCTGCT | ACACCACCTG | CAGTGTTAGC | 2820 |
| | TGCCACAAGT | AATGAACCGC | TTGTTGAAAT | TTGTTGTGCC | ACTGTCGCTT | GAATAGGTGC | 2880 |

| | TGGAGAGAAT | AATGGGAAAA | TTGCTCCCGC | TTTAGCAATA | CCTTGTCCAA | TTGCTACAGT | 3000 |
|----|------------|------------|------------|------------|------------|------------|------|
| | CAAACCACCG | TATGTCATAA | CTTTAGCAAT | AGCTAGGATA | GCTGAAATTG | TAAGGATCGG | 3060 |
| 5 | TAACCATAAT | TCTTTAATTG | CTTCGACCAA | TAAAGCACCT | GCACTTTTCC | ATTTTAACTT | 3120 |
| | CGTAATTAAA | ATTGTAATAA | TTACTGTTAA | TAAAATCGCT | GTCCCAGTTG | CACCAATTAA | 3180 |
| 10 | ATCGAGACGC | AACGCAATTC | CTTTAGGCGA | TAAATCACTC | ACAGTATTTG | GAATTGGCAA | 3240 |
| 10 | TTTTATTACT | AAACTTTCAA | GTGCACCTCC | AGGTTGGAAT | AATTTTTTGA | AGAATGGTGC | 3300 |
| | ACTCCATACT | AATACAAAGG | CAGTTAAAAT | TACGAACGGA | CTCCAAGCAA | AGACAATTTC | 3360 |
| 15 | TTTAGGCGTT | CGTTTTTGAA | TTTTATGTTC | AGACGCTTCC | AATCTGAAAA | TGTTTTTCGG | 3420 |
| | TTTAAATTTA | CGACAAACAA | ATGCTAACAC | CACCATTGTT | GCTAGTGATG | GAATAATGTC | 3480 |
| | TGCTAGTTCT | GGACCATGGA | ATATTGTTAA | TAATAATTGT | AATCCAGTAT | ATGTACCACT | 3540 |
| 20 | CACTGTTAAA | ATGACAGGTA | AAATTTCTTT | AATACCTTTC | ATACCATCTA | CAATGAATAC | 3600 |
| | TAAAACAAAT | GGAATAATAA | AGTTTAAAAT | TGGAAGTGTT | AATGCTGAGT | ATCTCGCAAC | 3660 |
| | ATCTAATGTT | GTAACGCCTC | CACTTAAGTT | AAACGTATCA | ATAATACTAA | CTGGTAAACC | 3720 |
| 25 | AATTGCACCA | AAGGCACCCG | CCGCACCATT | AGCAATTAAA | CATAACATCG | CTGCTTTTAA | 3780 |
| | TGGTTCAAAT | CCAAGTTGAA | TTAATAATAC | TGCACAAATC | GCAATTGGCA | CACCAAATCC | 3840 |
| 30 | TGCTGCACCT | TCTAAAAATG | CGTTGAAACA | AAATCCAATT | AATAATAGTT | GGATTCTTTG | 3900 |
| | GTCCACTGAA | ATACTTGCAA | TACTATCTTG | AATAATAGAA | AATTGTCCTG | TTTTAATAGA | 3960 |
| | AACTTTATAT | AACCAAACTG | CCATTAAAAC | GATATATCCT | ATTGGGAAAA | TACCGGCAAC | 4020 |
| 35 | AACGCCTTCT | GTAATCGCAC | CTGCTGATAC | ACGCGCTGGT | AATTCAAATA | CAAATAAAGC | 4080 |
| | CACAATCAAT | GTAACAACCA | AAGTTGTCAA | TGCTGCATAA | ATGCCTTTCA | TTTTAAAAAC | 4140 |
| | GGTTÄAGCAT | ATAAAAATA | AAATAATAGG | TACTGCTGCA | ACTAAGGCTG | ATAATCCGAC | 4200 |
| 40 | ATTATCGAAT | GGATTTACAG | TAAGTAGTGT | CATAATGACT | CCCTCTCTTT | ATATAAAATA | 4260 |
| | TTTATCATTC | TGATTAATCT | ACAACCTATT | TCAACTTATA | TTTTGCGATG | ATCACATATT | 4320 |
| 45 | TAAAATGTAA | CACTCCTATA | TGTGACAGGC | AATCGAATTT | TTACAAAAAG | TTCACAAAAT | 4380 |
| 40 | ATACACAATA | TTTAACTATA | ATAMATAATA | TATCATRETA | ATTATAAATA | CTAGATATTA | 4440 |
| | TTTATAATAA | TCTCAGGAAT | TCGCTTCAAA | ACTGCATCAT | GAGAGTTTAT | ATTTTTATTG | 4500 |
| 50 | AGAATCTCTC | ATTTTATGAA | TTGTAGGAAG | TAAACAAAAT | ATGACAAGCG | TCAAACCAAT | 4560 |
| | GATAATGATA | AATATCATAT | TAAACCATAG | TAAATTGAAT | TGATGATGGT | GTTGTATTTG | 4620 |
| | CCAAATTTCT | AATACTGTGA | AGATAGACAT | ATAGCTCATA | ATCTCTAAAT | TTAACGTACT | 4680 |

866

| AAATCGTTCA | TAGTATCTAC | CTGCAATGAA | AAATATAAGC | CAAATCACTA | TAAATGCGCT | 4800 |
|------------|------------|------------|------------|------------|------------|------|
| ATTAATCAAA | AGCAGCACCC | ATTTATCAGC | AAAATTATCA | GCATCCCCTG | СТАААТТАТА | 4860 |
| ATGAATAGGC | ACTTTGGTTG | GTAATTTTGG | ATAGGTCACT | ACTGTATAGC | ACATCATAGC | 4920 |
| TAAGTAAATA | AGTAGACTTA | ATATTGTAAA | AGACCTGATT | TTAGACATTC | TATCGCCTcT | 4980 |
| TCTTTACATT | TTATGTATAA | CACTCTGCCT | ATTTTACCTT | TTAATaCATT | ACCCCAAcGA | 5040 |
| TtAAaCAATA | tGTAaTGATA | CTATAATTGC | GTCAGGAGTA | TCCGCTTGTT | AAATGTGCAT | 5100 |
| AGCTTATATT | TAGCTGTTTA | ACATGCCACA | TAATGATTCG | AATTATT | | 5147 |
| | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 183:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1312 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 183:

| 60 | CATTATCTAA | nTTCATCAAA | AATGGATATA | ATAACTTTAA | CACCATTATC | CACTTACTTC |
|-----|------------|------------|-------------------|------------|------------|--------------|
| 120 | GATTACTTCT | ACAACTCTGC | CAATTAACTT | CACCATCCAA | CCTACACCAA | AGGCGTCGCA |
| 180 | AAAAGATGCA | CAACCGCTTC | TCAGGTAGTC | CGTAATACAA | CTTTCACnTG | TCAGCAGCAA |
| 240 | TCTAATTTGT | CTTGAATCTG | TTAATATGTG | ATATGTTTGT | TAAGTCGTGG | CCAGTTACTC |
| 300 | AATTGTAAAT | AACGATTGAC | CTTTTCGGCA | CTGTGGCATA | CGACATGGTA | TGAATATGAC |
| 360 | CGATACTAAT | TACGTACAAT | CTTAAATCTC | CATCATTTGA | CTTTAAATGT | TCAGGATCAC |
| 420 | CGCACGAATC | TACCTACATA | TCACCTGGTT | AACCACAGTA | TATGATCATC | TCATTATCTG |
| 480 | ACATGCGGTA | ATGTCCAAGT | CATTTTTTCG | AGTAAATGGC | CTTCCGGTGT | AAAACCTTAC |
| 540 | ATTITCAATG | CATCATGGGT | AAGCCAGTAC | CGCAATTACG | CACTCGTTCT | ATGTCTGTAT |
| 600 | CGTTTTAAAG | TACTATCCAT | GCAACAGTCG | TACAACAGTT | CAAATGCCAA | TCTTTTTCAT |
| 660 | TGTCGTTACT | CTTGATGTGG | TTTAAAAACA | TCCGAACCAA | CTGGATCTTG | TAATCAAATG |
| 720 | TTGTTTTTGA | CAATTTTATA | TCATTGTAAA | TTCTAGTTGA | CATACACTTC | AATACGCCAT |
| 780 | ATTTTTAACA | TCACACCTTT | TAGCGTATTG | TGACGTATTG | TATCATCCAC | GATGTAATTA |
| 840 | GAATTGTCCT | TAAATTGTTT | AAACCATGCT | AAATGAGCTT | ATGCTTCAAT | TCTTGTTCTA |
| 900 | ATCCTTCATA | TATTTTTCTC | TTCAGACGCT | TTGTCTTTGT | CAGGATATAA | TTCGGTGCGC |
| 960 | CCTACTCATC | DATTACCERA | لا لايلملململململ | TCCTTTCT | CACTTCCCAA | CCALALALVACA |

55

5

10

15

20

25

30

35

40

45

| TCAAGTACCT | CATTACCTAA | TCTTGCTCTG | AAAAATGCAC | CAACAGAAAT | GTCACCATCC | 1080 |
|------------|------------|------------|------------|------------|------------|------|
| TGCATTTGAG | TAGGTTTTTT | TAATAAATCA | AACCCTGCTC | TTAATTTACC | AAGTGGCGAT | 1140 |
| ATTAATTTTG | TAGTAACAAA | TGGTTTAATA | TCTGTTGGAA | TACCCATAAT | TGAACCACCT | 1200 |
| GGAATCGGAT | ATAATTTATT | TTTCGCAAAA | ATATATGATT | GTCCAGTCGT | ATTTGTAACA | 1260 |
| ATATCTTGTT | CTAATCCAAT | ATCTTTCGCT | AATTCTGTCA | TAATCGTTTT | TC | 1312 |
| | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 184:

5

10

· 15

25

30

35

40

45

50

55

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6157 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 184:

| 60 | ACAATTTATG | GGAACGGAGG | TCGTATATAA | ATACTACTTG | AAAATATGAT | TTTTACAATA |
|-----|------------|------------|------------|------------|------------|------------|
| 120 | TGTTGTACTA | CATTAATAAC | GATTGTATTG | ATTAATCATT | TAATCGTATT | CATACATTTT |
| 180 | TGAGCAGTTA | GTGGTGGTGC | GGTGCCATCA | TGGACTTTCA | GTAAAAGCAG | CTCCAAGAAG |
| 240 | AATTATTTA | ATAGATTAAC | TTATTCTTAA | TGGCGTCGAT | AAAAACAACG | TTCGGTAAAC |
| 300 | AGGTCCGGCG | TTGGTATGTA | ATAAGTTATC | TATGATTTGC | TTTTTGTACT | TCAATATTAT |
| 360 | AATAAGCTAT | ATAGTTTAAC | TAAGAATGTA | TATTTATAAT | CGGGCTTTTT | ATGTAAATGT |
| 420 | AGATAAAATT | AGATTTATGC | AGGGAGCGTT | AAGTATGCAA | TAGCCTAGTT | GTAAAATATA |
| 480 | ATGGTTTTAC | TTATTACTAC | ACGTGCCGTG | AGGAAGGTAA | TTCTTTTTTG | ACCAAAACCT |
| 540 | GTTATACATC | CAAAAGAAAG | TCGATTTTTA | GTCAATTAGG | TCTGATGTTC | AGGCAATTCG |
| 600 | AATCTAGTCC | GAAATACTGA | ACCACCAGAT | GCCACGCGGC | CAATATGAAG | ATATGCACCG |
| 660 | AAGGTTATGA | CTTGTTGAAC | TTATGATTAT | CGTTAGATGG | TTTAAAGATG | TTTCGTTTGG |
| 720 | TAAGCTTAAA | GCTTTAAAAT | TGGGGATTTT | TATCATTAGG | GTTGCTGGTC | TGAAATTGTT |
| 780 | CTGAAGGTGC | GGTGGCAAAA | TGCTCCTATG | TAACGATGTG | AAGGGTATTG | TAGAGATGTA |
| 840 | GTAAAGATCA | AAGTATGAAG | CAATTTTAAA | AATATGCACG | GGCTTTTTAG | CATTTATGAA |
| 900 | AAGAACTAAG | GAAACTTTAA | TAAACCAACT | TGGATCATTT | GATAATGAAA | AGAGACTATT |
| 96 | TTTTAGTGAT | TTGGATCCTA | TGATGAAGTG | AAGAGCAAGT | GATACGATTA | TGAAGCATTA |
| 102 | ATGACCATGT | AATTATATAT | ACAATCCGCA | TGATTGATCC | AACGACAATA | TCAAGCAGAA |
| 108 | TTACGATTGA | GGACATGTTA | CAGTGAATCT | TCAAGTGGTA | GACAAAAATA | AGATTCTGAT |

| | AGAATAAAA | GAGATTTTAA | CATTAGAAAG | GAGGGGCATA | ATGAATTTAA | AGCAATCTAT | 1200 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | AGAAGAGATT | ATTAATCAAC | CTGAATATGA | ACCTATGTCA | GTGTCAGATT | TTCAAGATGC | 1260 |
| 5 | ATTAGGTTTA | AGCAGTGCCG | ACTCGTTTAG | AGATTTAATT | AAGGTGCTTG | TGGAGTTAGA | 1320 |
| | ACAATCAGGA | TTAATCGAAC | GTACAAAAAC | AGACAGATAC | CAAAAAAAGC | ATAGTTATAG | 1380 |
| | AGGTCAATCA | AAATTGATAA | AAGGAACGTT | AAGTCAAAAT | AAAAAAGGCT | TTGCATTCTT | 1440 |
| 10 | AAGACCTGAA | GATGAGGATA | TGGAAGATAT | ATTTATTCCC | CCGACGAAAA | TTAATCGTGC | 1500 |
| | CTTGGATGGA | GATACTGTTA | TTGTAGAAAT | CCATCAATCA | AAAGGTGAAC | ATAAAGGTAA | 1560 |
| 15 | AATCGAAGGG | GAAGTTAAGT | CGATTGAGAA | GCATTCTGTA | ACTCAAGTTG | TTGGTACGTA | 1620 |
| | TAGTGAAGCT | AGACATTTTG | GCTTTGTTAT | TCCGGATGAT | AAACGTATTA | TGCAAGATAT | 1680 |
| | TTTCATTCCT | AAAGGTCAAA | GTTTAGGCGC | AGTCGATGGT | CATAAGGTAC | TTGTACAAAT | 1740 |
| 20 | TACTAAGTAT | GCTGATGGTT | CAGATAATCC | AGAAGGACAT | ATTTCTGCTA | TTTTAGGACA | 1800 |
| | TAAAAATGAT | CCTGGCGTAG | ATATTTTATC | TATTATCTAT | CAACATGGCA | TAGAAATTGA | 1860 |
| | ATTTCCTGAT | GAAGTGTTAC | AAGAAGCTGA | AGCAGTACCT | GATCATATTG | AAAATACTGA | 19,20 |
| 25 | AATTAAAGGC | CGTCATGATT | TACGTGATGA | ATTGACAATC | ACAATTGATG | GTGCTGATGC | 1980 |
| | TAAAGACTTA | GATGACGCAA | TTAGTGTTAA | AAAGTTAGCG | AACGGTAATA | CGCAATTAAC | 2040 |
| 30 | TGTAAGTATT | GCTGATGTCA | GCTATTATGT | AACAGAAGGT | TCTGCATTGG | ATAAAGAGGC | 2100 |
| | ATATGATAĞA | GCGACAAGTG | TATATCTTGT | TGACCGTGTA | ATTCCAATGA | TTCCACATCG | 2160 |
| | ATTAAGTAAT | GGTATTTGTT | CATTGAATCC | TAATGTTGAT | CGTTTAACTC | TAAGCTGTCG | 2220 |
| 35 | CATGGAAATC | GATGCTAGTG | GTCGCGTTGT | TAAACATGAA | ATTTTTGATA | GTGTTATACA | 2280 |
| | TTCTGATTAT | CGAATGACGT | ATGATGCGGT | AAATCAGATT | ATTACTGAAA | AGGATCCTAA | 2340 |
| | CATTCGCGAA | CAATATAATG | AAATTACGCC | TATGCTAGAT | TTAGCACAAG | ATTTATCTAA | 2400 |
| 40 | TCGTTTGATT | CAAATGAGAA | AACGACGTGG | TGAAATCGAT | TTTGATATTA | GTGAAGCAAA | 2460 |
| | AGTATTAGTT | AACGAAGACG | GTATACCAAC | AGATGTTCAA | TTAAGACAAC | GTGGCGAGGG | 2520 |
| 45 | TGAACGTCTA | ATTGAATCAT | TTATGTTAAT | TGCAAATGAA | ACAGTTGCTG | AACATTTTAG | 2580 |
| 45 | TAAGTTAGAT | GTACCTTTTA | TTTACCGAGT | GCATGAGCAA | CCTAAATCAG | ATCGCTTAAG | 2640 |
| | ACAATTCTTT | GATTTTATTA | CAAACTTTGG | CATCATGATT | AAGGGTACTG | GCGAAGATAT | 2700 |
| 50 | TCATCCAACA | ACACTTCAAA | AGGTTCAAGA | AGAAGTAGAA | GGTCGACCTG | AACAAATGGT | 2760 |
| | CATTTCAACA | ATGATGTTGC | GTTCAATGCA | ACAAGCGCAT | TATGATGATG | TGAACTTGGG | 2820 |
| | ACATTTTGGC | TTATCAGCTG | AATATTATAC | GCATTTTACA | TCACCAATTA | GACGTTATCC | 2880 |

| | AGAAGTGAAG | CGTTGGGAAG | ACAAATTGCC | TGAGTTAGCT | GAACATACTT | CTAAACGTGA | 3000 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | ACGTCGTGCT | ATTGAGGCAG | AACGTGATAC | TGATGaATTG | AAAAAAGCAG | AATATATGAT | 3060 |
| 5 | TCAACATATT | GGTGATGAAT | TTGAAGGTAT | TGTCAGCTCA | GTAGCTAACT | TCGGTATGTT | 3120 |
| | CATTGAATTG | CCAAATACGA | TAGAAGGTAT | GGTTCATATT | GCGAATATGA | CTGATGATTA | 3180 |
| | TTACCGTTTT | GAAGAGCGTC | AAATGGCATT | AATTGGTGAG | CGTCAAGCTA | AAGTATTTAG | 3240 |
| 10 | AATTGGTGAC | ACAGTTAAGG | TTAAAGTGAC | GCATGTTGAT | GTAGATGAAC | GATTAATTGA | 3300 |
| | TTTTCAAATT | GTAGGTATGC | CTTTACCGAA | AAATGATCGA | TCACAGCGCC | CAGCGCGAGG | 3360 |
| 15 | TAAGACAATT | CAAGCCAAAA | CGCGTGGTAA | ATCATTAGAT | AAATCAAAAT | CTGATGATAA | 3420 |
| | GGGTCGTAAG | AAAAAAGGTA | AGCAACGTAA | AGGTAAAAAC | CAACGTAATA | ATGATAAATC | 3480 |
| | AGGTAATAGT | AAGCATAAGC | CATTTTATAA | AGATAAAAGT | GTGAAAAAGA | AAGCACGTCG | 3540 |
| 20 | TAAGAAAAA | TAAGCAGCAA | TGAGGTGAGT | ATGAATGGCT | AAGAAGAAAT | CACCAGGTAC | 3600 |
| | ATTAGCGGAA | AATCGTAAGG | CAAGACATGA | TTATAATATT | GAAGATACGA | TTGAAGCGGG | 3660 |
| | AATTGTATTG | CAAGGCACAG | AAATAAAATC | AATTCGCCGA | GGTAGTGCTA | ACCTTAAAGA | 3720 |
| 25 | TAGTTATGCG | CAAGTTAAAA | ACGGTGAAAT | GTATTTGAAT | AATATGCATA | TAGCACCATA | 3780 |
| | CGAAGAAGGG | AATCGTTTTA | ATCACGATCC | TCTTCGTTCT | CGAAAATTAT | TATTGCACAA | 3840 |
| 30 | GCGTGAAATC | ATTAAATTGG | GTGATCAAAC | ACGTGAGATT | GGTTATTCGA | TTGTGCCGTT | 3900 |
| 50 | AAAGCTTTAT | TTGAAGCATG | GACATTGTAA | AGTATTACTT | GGTGTtGCAC | GAGGTAAGAA | 3960 |
| | AAAATATGAT | AAACGTCAAG | CTTTGAAAGA | AAAAGCAGTC | AAACGAGATG | TTGCGCGCGA | 4020 |
| 35 | TATGAAAGCC | CGTTATTAAG | CGATTTAGTT | GCTTAATCGG | GCTATATTTG | ATATAGTTAT | 4080 |
| | ATGTGCTTTT | GTAAATTACA | AAAGTATGAT | TTGTTTGATT | TATTATTTCG | GGGACGTTCA | 4140 |
| | TGGATTCGAC | AGGGGTCCCC | CGAGCTCATT | AAGCGTGTCG | GAGGGTTGTC | TTCGTCATCA | 4200 |
| 40 | ACACACACAG | TTTATAATAA | CTGGCAAATC | AAACAATAAT | TTCGCAGTAG | CTGCCTAATC | 4260 |
| | GCACTCTGCA | TCGCCTAACA | GCATTTCCTA | TGTGCTGTTA | ACGCGATTCA | ACCTTAATAG | 4320 |
| | GATATGCTAA | ACACTGCCGT | TTGAAGTCTG | TTTAGAAGAA | ACTTAATCAA | ACTAGCATCA | 4380 |
| 45 | TGTTGGTTGT | TTATCACTTT | TCATGATGCG | AAACCTATCG | ATAAACTACA | CACGTAGAAA | 4440 |
| | GATGTGTATC | AGGACCTTTG | GACGCGGGTT | CAAATCCCGC | CGTCTCCATA | TITGTAGCCT | 4500 |
| 5 0 | ACAGCCTTTG | TGGTTGTGGG | CTTTTTTATT | TTGTGTTTTT | CAGGGGATAA | TGCATTGCAG | 4560 |
| | AATTTGTTGT | GAGTATTGAT | ATAGCAGTGT | TTGTATAGGT | GTTTATTTGA | TGGAGGAAAG | 4620 |
| | AGTAATAAGT | GATTATGAAT | TAGTTTTTGA | GATATAAGGG | GACAGTGATG | TGTGTCAAAT | 4680 |

870

| | TTATACGCAA | AAAATTCTCC | ATGTTATATA | TGTCAATATA | AAAATGTGAA | TCGTCTACAC | 4800 |
|----|------------|------------|------------|------------|------------|------------|---------------|
| | TTAATTGGAT | AAATGGCTAC | TGAAAAAGAA | CTTTTCATTT | TTGTTACGTC | ACTAAGTGGG | 4860 |
| 5 | TGTAGTTATA | AAGAGATGAG | CCGAGTTTTG | ATATTTTCAT | TAGAATCAAT | ATGCCTATTA | 4920 |
| | ACACAATCAG | CAATAGTTGA | CGAGACGGAA | ATAAAAGAAG | TCGTAGTTAA | GAAATGCATT | 4980 |
| 10 | TCACAACATA | CCATTGTAGC | CATTTTTATT | GTTTTGGATG | ATAAACTCTT | TTTGGAATTT | 5040 |
| 10 | TTAGTTTTTA | TAATTTGCAA | CTACACTACT | TCTTTTACTA | ATATTAATGT | CTAAGTAATC | 5100 |
| | GATAAAAAAT | TTTCCATTGA | ATAAATGAGA | AGTTAAAAAC | TTTACTTAAC | CTTTCycATT | 5160 |
| 15 | GCATTTTCCT | ATTCACGATT | TTAAGAACCC | AACATACTAC | AAACGAATTT | TAAAAGGCGA | 5220 |
| • | GAGTAAAGCT | TACTTGTTTA | TTATACATAT | TTAAAATCCA | AGAGTCAGAA | CAGACTACTC | 5280 |
| | CTCTTTATAA | СТАТАААААА | TAGCTATGAA | AAAATCTATC | GTCATAGATT | CCTTCATAGC | 5340 |
| 20 | TAATCTTAGT | ATGTTTATTT | TTATTTTAGG | ATGCTATTTA | TCAACTCAAC | ATATAACTCA | 5400 |
| | CTATTTTTAT | AACCTTCTAA | TATATCATTA | ACTTGTCTAA | TAGGTATTTC | TGGTACTTCT | 5460 |
| | CTAATGTTTT | CCAATTTTGT | TTTAAATTGT | TTTTTTGTTA | TTTGCTCTTT | ATTTGTAGCC | 5520 |
| 25 | AATTGGAACA | AGTAAGAATC | TAGCATATTA | ATTTCTTTAT | ATGAATACAT | ATATCTTAAT | 5580 |
| | AACACTAAAT | CTCTAGTTTT | TAAGTTAGGC | GCTAGTTCTT | CTTGTAATTG | TTCTATTGAT | 5640 |
| 30 | TGTYTCATTA | ATAACAATCT | CATTTCTAAT | TCTTCATTAT | TCATTTTATC | ACACTCTTLT | 5700 |
| | TATATTAATG | CTTGACCAAC | TTGGGAAACC | CAAAACCCTA | TGCTTCTTGC | AGTAGAATCT | 5760 |
| | TTAATACCAG | TTCCCATCAA | TGCTTGTGAA | ACTTGACCTT | GTACATTTCC | CCATGTAGCC | 58 <u>2</u> 0 |
| 35 | TCTTCTTGTT | TTAATGCATT | ATTCAATGCG | GGATTTACAA | ATTTATCCCA | TCTTTTTTT | 5880 |
| | ATGATTTTCC | GGCACGGGGA | CTGATTTCTT | TAACACCATT | AAACACAGAT | TTTTTATTTT | 5940 |
| | TAATCATAGC | TTTATAGTAT | CATGTTGGCT | AAGCTATAAA | TAAGTCAGTT | TCTCTAAAAA | 6000 |
| 40 | TTAAATAACT | GAATGTAAGA | CAATCAACAA | WCCAAATTTA | TACTTCATCT | AAACCACTGT | 6060 |
| | GGTCGTCATC | TTTTTGCTTT | TCTTTTTCTT | TCTCTCGTTC | TIGTICITIT | TTGTACTCTT | 612 |
| 45 | CTTCAAATTC | TTTTTCTTTC | TTTTCTACTT | CTTCTCT | | | 615 |
| 70 | | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 185:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 884 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

| | (2) INFORMA | TION FOR SE | O TO NO. 18 | ı | | | |
|----|-------------|-------------|-------------|------------|------------|------------|-----|
| | TTGAAAGCGC | TTTACCAACC | TGTTAATATA | TAATAGTAAT | ATAC | | 884 |
| 25 | TACYTTTAAG | TTTGACGAAA | AAAATGGGTT | CAAnaaacac | GTAGTATTTA | TGnaaaagta | 840 |
| | AGTATTTAAG | TAAAGCYTTG | AATTTGATTT | CyGaGTTACC | aGATGGACaT | CCGaGATCAC | 780 |
| | AATGTATCCA | AATCATTAGA | AAATCTGACA | GCATCGATGA | GGCTAAGGCA | GTAAGTTCGA | 720 |
| 20 | CAGACTTCAA | ATTGAAAATC | GAACAGTTAC | GTCGTGATAG | TGAACGCAAA | GAATTTGAAG | 660 |
| | GAAGTGATTT | GCTTAATGGT | CATATTACGT | TACCGATETT | ATTAGAAATG | CGTAAAAATC | 600 |
| | AAATCATTGA | TGATGTATTA | GACTTCACAA | GTACCGAAAA | GAAATTAGGT | AAGCCGGTCG | 540 |
| 15 | CTGATAAAGA | GACTGTACGA | AAATTGAAAA | TGATTGGTCA | TTATATAGGT | ATGAGCTTCC | 480 |
| | ATCGCAAAAC | AGCACTGTTA | ATTCAAATAT | CAACTGAAGT | TGGTGCAATT | ACTTCTCAAT | 420 |
| | TCCAATTTCA | AGACCAATTT | AACAGTCAAC | AGACAATTAT | TAATTATTTA | CGACGTATCA | 360 |
| | ATAATCGTGT | ACATCAATTG | ATATCTGAAT | CTATCGTTGA | TGTTTGTAGA | GGGGAACTTT | 300 |
| | CTATTTTAAC | TGGGAATTTT | TTATTGGCAT | TAGGACTTGA | ACACTTAATG | GCCGTTAAAG | 240 |
| 5 | ATAAAAGCGA | CAAGCGTCGA | GGCAAGTTAA | CCATATCAAA | GAAATGGGAT | CAGACAACTG | 180 |
| | AAGTTGCAGT | CGCATTAGAG | TTAATTCATA | TGGCAACACT | TGTTCATGAT | GACGTTATTG | 120 |
| | CATTTGTTAT | TCTGAGTAGC | CAATTTGGCA | AAGATGAACA | AACGTCTGAA | CAAACGTATC | 60 |

30

35

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6876 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 186:

| 40 | AATTTCATCT | GCTCGTGCAA | AATCTTTGTT | TTTCCTTGCT | TCATTACGCT | CTTCGATTAA | 60 |
|----|------------|------------|------------|------------|-------------------|------------|-----|
| | TTTTTCAACA | TCTTCATCCA | ATAATTCATC | TGCATTTTTA | GATTTTAACG | GTACACCTAA | 120 |
| | AACATCGCTG | AAAATTTGAT | AAACTGCTTT | AAATTTATCA | ATTACTTCTG | TTGATGTTGT | 180 |
| 45 | GTTCTCTAGT | ACATATTTAT | TCGCAAGTKT | TGCTAAATCA | TACCAAGCTG | TAATTGCATT | 240 |
| | AGCTGTATTA | AAATCATCAT | TCATAACTGT | TTCAAAACGA | TTTAAAATCG | CATCAATTTG | 300 |
| 50 | ATCAATATAT | GTCTGTTGAT | TTTCAATATT | AGTAGCAATT | TGTGCGCGCT | CTTCAATTAA | 360 |
| | TTGATAACTA | TTGCGAATAC | GCTCTAGTcC | aCTACGTGCT | GATTCTACCA | ATTCTAGATT | 420 |
| | ATAGTTAATT | GGGCTTCTAT | AATGTACGCT | AATCATAAAG | א אייריידיא מיייא | CATCTCCATC | 400 |

| | ATTATCAATA | TTAATGAAAC | CATTATGCAT | CCAATAATTA | GCAAATGGCG | CATGATTATG | 600 |
|------|------------|------------|------------|------------|------------|------------|------|
| | TGCTTCTGAT | TGTGCTATTT | CATTTTCATG | ATGTGGAAAT | TGTAAATCTG | AACCACCCGC | 660 |
| 5 | ATGTATATCA | ATTGTAGGTC | CTAGCTCATG | AAATGCCATT | ACAGAACATT | CTATATGCCA | 720 |
| | TCCTGGTCTA | CCTTCACCAA | ATGGGCTATC | CCAACTAATC | TCGCCAGGTt | CGCTTTTTTC | 780 |
| o | CACAATGTAA | AATCAAGTGC | ATCTTCTTTA | TGCTCTCCTG | CATCTATACG | AGCACCCACT | 840 |
| J | TTTAAGTCAT | CTATGGATTG | ATGACTTAAT | TTACCATAAC | CTTCAAATTT | ACGTGTTCTA | 900 |
| | AAGTAAACAT | CGCCACCACT | TTCATATGCA | TAACCTTGAT | CCACCAAATC | TTTAATAAAT | 960 |
| 5 | TGAATAATGT | CATCCATATG | GTCCATTACC | CTTGGATTTG | AAGTCGCTTT | TCTAACATTT | 1020 |
| | AACGCACCAA | CATCTTCATG | AAAAGCAGCG | ATATATTTT | CTGCAATTTC | GGGAACAGAC | 1080 |
| | TGATTTAATT | CTTGAGAACG | TTTAATTAAT | TTATCATCTA | CGTCTGTAAA | ATTTGATACA | 1140 |
| 20 | TATTCTACAT | TATATCCTTG | GTATTCAAAG | TAACGTCTCA | CTACGTCATA | ATTAATTGCW | 1200 |
| | GGTCTTGCGT | TACCAATATG | AATGTAGTTA | TATACAGTAG | GACCACATAC | ATACATTTTT | 1260 |
| | ACTTTCCCTG | GTTCTATAGG | CTTGAACACT | TCTTTTTGAC | GTGTAAGCGT | ATTATATAAT | 1320 |
| 25 | GTAATCATCT | TGAATCTCTC | CATTCCTAGT | CTTTTCAAGT | TGTCGTTCTA | AATGCTTAAT | 1380 |
| | TTGTTCATAA | ATTGGATCAG | GTAGATGGCG | ATGATCAAAT | GTTTTTCCAA | CTCGAACACC | 1440 |
| | ATCTTGCTTA | ACAATATGTC | CTGGTATACC | AACAACCGTT | GAATAACTTG | GAACTGATTG | 1500 |
| ,, | TAAAACAACT | GAATTTGCAC | CAATATTTAC | ATTTGAATTT | ATTTTAATAT | TTCCTAAAAC | 1560 |
| | TTTCGCACCG | GCTGCTATTA | AAACATTGTC | TCCTATATCT | GGGTGTCTTT | TCCCTCTTTC | 1620 |
| 35 | TTTCCCTGTC | CCACCAAGTG | TCACGCCTTG | ATAGATTGTC | ACATTATCAC | CAATTGTACA | 1680 |
| | TGTTTCTCCT | ATTACAACGC | CCATACCATG | ATCTATAAAT | AGACGCTTTC | CAATTTTAGC | 1740 |
| | ACCTGGATGG | ATTTCTATAC | CTGTGAAAAA | TCTTGAAATT | TGAGATATCG | CGCGTGCTGC | 1800 |
| 10 | AACATATTTT | TTTTGGTTGT | ATAACTTATG | TGCAATCAAA | TGACTCCAAA | CTGCATGTAA | 1860 |
| | ACCTGCATAC | GTTGTAATGA | CTTCTAATGT | TGAACGTGCC | GCTGGATCCT | GCTCAAATAC | 1920 |
| | CATTTTTATA | TCGTCTCTCA | TTCTTTTTAA | CAAGATCATT | TCCTCCTCAA | TGATTGAACT | 1980 |
| 15 | ACGTAAATAC | ATAATTGAAG | TACCTGCGAA | ATTAAATATC | AAAAAAGCAC | CACTAACATA | 2040 |
| | CAAATTGTAT | TGTTAGAGGC | GCTTCCGCAC | GGTTCCACTC | TGAATTTAGO | GAATAACATT | 2100 |
| F.O. | AATAATATTG | CGGCCCTTC | CAAATTATCA | AGGAAACTAA | GTCAACTTAA | TGCTCATCAC | 2160 |
| 50 | TCTCATTATA | TATTTAATTC | ATTTTACGAA | GGTGCATTCA | TTAATTICTA | CGTTGTACTC | 2220 |
| | ACAGCAACCG | TACACTCTCT | GCATCGTATA | ATTTAATTA | CTAATCCTTC | GTTTTATATA | 228 |

| | ATAAAATTCA | AGTATATACT | ACCTTGATCT | TGTCTATTTC | ATTACTTATA | TTGTTTTAAA | 2400 |
|-----|------------|------------|------------|------------|------------|------------|------|
| | CGGTTTAGCA | CTTTTTCTTT | ACCAAGTACT | TCAATTGTAT | TTGGTAATTC | AGGACCATGC | 2460 |
| 5 | ATTTGGCCTG | TTACAGCAAC | ACGAATAGGC | ATAAATAATT | GCTTGCCTTT | TATTCCTGTT | 2520 |
| | TCTTTTTGAA | CTTCTTTAAT | TGTCTTTTTA | ATTTCAGCCG | CTTCAAATGG | TTCAAGTGCT | 2580 |
| 10 | TCTAATTTAC | TGAATAAGTG | CGTCATTAAC | TCTGGTACTT | GCTCTCCATT | AATCACTTGT | 2640 |
| ,,, | TGTTCTTCTT | CACCAAGAGC | TGGCATTTCT | TTAAAGAACA | TTTCTGATAA | AGGTACAATT | 2700 |
| | TCACCGGCAT | AACTCATTTC | TTTTTGATAA | AGCGCAATTA | ATTTGCGTCC | CCAAGATAAA | 2760 |
| 15 | TCCTCTTCTG | ACGGCACCTC | AGGAATCAAA | TTTGCTTTAA | TTAAATGAGG | TAATGCTAAT | 2820 |
| | TGGAATACTG | TTTCAGTATC | TTTTTGTTTC | ATATATTGGT | TATTAACCCA | TGCTAATTTT | 2880 |
| | TGCTTATCGA | AAAATGCTGG | TGATTTTGAC | AAACGCTTTT | CATCAAAGAT | TTTGATAAAT | 2940 |
| 20 | TCTTCTTTAG | AAAAGATTTC | TTCTTCACCT | TCAGGAGACC | AACCTAATAA | CGCAATAAAA | 3000 |
| | TTAAATAACG | CTTCAGGTAA | ATAACCTAAG | TCACGATATT | GCTCAATAAA | TTGTAAAATT | 3060 |
| | TGCCCATCAC | GTTTACTTAA | CTTTTTACGT | TCTTCATTAA | CAATTAATGA | CATATGACCA | 3120 |
| 25 | AAACGAGGTG | GCTCCCAGCC | AAATGCTTCA | TAAATCATAA | TTTGTTTAGG | CGTGTTTGAA | 3180 |
| | ATATGATCAT | CACCACGAAT | TACATCTGAA | ATTTGCATGT | AATGATCATC | TATAGCTACT | 3240 |
| 30 | GCAAAATTGT | ACGTTGGAAT | GCCATCTTTT | TTTACGATAA | CCCAGTCACC | AATACCATTT | 3300 |
| | GAATCAAATG | AAATATTTCC | TTTTACCATA | TCATCAAATG | AATACGTTTG | GTTTTGAGGT | 3360 |
| | ACTCGGAAAC | GAATTGATGG | TTGGCGTCCT | TCTGCTTCAA | ATTGTTGACG | TTGTTCTTCA | 3420 |
| 35 | GTCAAATGCG | CATGTTGACC | ACCATAGCGA | GGCATTTCAC | CACGAGCGAT | TTGCGCTTCA | 3480 |
| | CGTTCAGCTT | CTAATTCTTC | TTCTGTCATA | TAGCATTTAT | ATGCTTTATC | TTCTGCTAGT | 3540 |
| | AACTGATCTA | TTAATGGTTG | GTAGATATGT | TGACGTTCAG | ATTGACGATA | TGGTCCGTAG | 3600 |
| 10 | CCATTGTCTT | TATCTACAGA | CTCATCCCAA | TCTAATCCTA | ACCATTTAAG | ATTATCAAAT | 3666 |
| | TGTGATGTTT | CTCCATCTTC | TAAATTACGT | TTTTTATCAG | TATCTTCAAT | TCGAATCACA | 3720 |
| | AAATCTCCGT | TGTAATGTTT | AGCATACAAG | TAATTGAATA | ATGCTGTTCT | TGCATTACCA | 3780 |
| 15 | ATATGAAGAT | ACCCAGTTGG | ACTTGGTGCA | TATCTTACTC | TTATACGATC | GCTCATTTTT | 3840 |
| | TTCACTCCTA | AATTAAATAT | CAGATTTTCA | AGTTAGTTCA | TATAAATTGT | TCATTTGCTA | 390 |
| 50 | TCTTCGACCG | TCATAACAAA | TGTCTAACTC | GTCTTATTGT | TAAAACGAAA | CAATGCTTTT | 396 |
| | TAACATGACC | TTAAAATAAT | TTCATTGTTT | AATCATAACA | TAATTCCCTG | GGTAATATGC | 402 |
| | TTAAATTTTA | AATAGAAAGC | TGTTGTTTTT | TCAACACTTT | AAAAAAGCTA | TCCCTAAGAA | 408 |

| | TTAAACTTCA | AATTAACTAT | TCAAATACGT | TAAAATTGAT | TCTAATTTTG | TATGTCTTGA | 4200 |
|----------------|------------|-------------------|------------|------------|------------|------------|----------|
| | TTGCTATAAG | AATAACTTTA | TTAATATCTA | AAATTTAACA | CTTAATGAAC | TTGTTTCAAT | 4260 |
| 5 | GATATATTAG | CACTATTTGT | ATTTTTTGAT | AACTAATATG | TTTTGCATTT | ATTTATAGTT | 4320 |
| | ATACTTCAAA | TTACAAACTt | CGCCATTTCA | TATACCTTTT | AATATCTATT | TTGTTTTCGT | 4380 |
| 10 | CAACTACAGT | TTTTATAATG | ATACTGTATC | TTCGATTTTT | TTAGCAAAAA | CAATTCTTCC | 4440 |
| | TGAAGATGTT | TGCAATAAGC | TGACTACTTC | TAAATTGACA | TGACTGCCAA | TAAGATTTTT | 4500 |
| | AGCATTATCA | ACAACTACCA | TCGTACCATC | ATCTAGATAT | CCTACTGCCT | GACCAGGCtC | 4560 |
| 15 | CTTACCCATT | TTTGTCAGTA | AAATATGCAG | TTGATCACCT | TGATGTACAT | TAGGTTTGAT | 4620 |
| | TGCTTCTGAT | AAATCATTAA | CATTTAATGC | TTTGATACCA | TGTACATGAC | AAACTTTATT | 4680 |
| | TAGGTTGAAA | TCTGTCGTTA | TAATACTTGC | ATGATATTGT | TTTGCTAATT | TTAATAACAT | 4740 |
| 20 | CGTATCAATA | TCACTATGTG | TTTTAGTTGG | ATGTATAACC | TTTGTAGGAT | AGTCTAAATC | 4800 |
| | ATACAATTCA | TTTAAAATAT | CTAAGCCTCT | TTTACCCTTT | TCaCGTTTAA | CACTGTCATT | 4860 |
| | TGAATCTGCA | ACAATTTGTA | ATTCATTAAT | AACACCTTGT | GGAATTAAAA | TATTGCCATC | 4920 |
| 25 | GATAAAACCG | CAACGAATGA | CTTCTAAAAT | ACGACCATCA | ATAATTGCGC | TTGTGTCGAT | 4980 |
| | AATTTTTGGC | GTAgcaCTTT | TAGTATGTTG | TGACATGGAA | CGCGCTATAT | TCTCAGGTAA | 5040 |
| 30 | AAACATTAAC | ATTTCATCTC | GTTTTTTAAG | GCCAAATTGG | AAACCGAAAT | AACATAGTAA | 5100 |
| | TATCGTAATT | ATGACAGGAA | TGAAATGATT | AAAAATAGAG | TTGCCAATTG | ATTCTAATAT | 5160 |
| | AAACGACACC | ATAACAGAAA | TAAGTAATCC | GATTATTAAA | CCTATTGTTG | CGAATAGTAT | 5220 |
| 35 | TTCAACAGCA | CTTCTACGCA | TAATAAAATG | TTCTAAACCT | TTTATAGCGT | TAGTAACTCG | 5280 |
| | TCTAATAAAT | ACACCAAAAA | TTAAGAACAT | AAAAATACTA | CCGATAATGC | CATCTACATA | 5340 |
| | GTGATTTTT | AAAAAGCTGG | AGTTTTGTAA | TCCAAGATCA | TTTGCAATTT | CAGGAATAAT | 5400 |
| 40 | AATTATTCCT | AATGCGCTCC | CAATAATTAA | GTAAATAATA | ATAACCATTA | GTTTAACGAT | 5460 |
| | ATTCACACAA | TGTCCTCCTT | TCTTGATGTT | TTATGAATGA | AGAGCAAATG | ACAATACTTC | 5520 |
| | ATGTACAGTA | GTTACACCTA | TTACTTGTAT | ACCITCAGGA | TATGTCCATC | CGCCTATATT | 5580 |
| 45 | ATTTTTAGGA | ATAATTACAC | GTTTGAAACC | TAGTTTTGCA | GCCTCTTGCA | CGCGTTGTTC | 5640 |
| | TATCCGAGAT | ACACGACGTA | CCTCACCCGT | TAAACCAACT | TCTCCAATAT | AGCAATCTAA | 5700 |
| 50 | TCCGTCGACA | GCTTTÄTCTT | TAAAGCTAGA | TGCAGTTGCT | ACAATTACAC | TTAAATCAAC | 576 |
| = - | TGCTGGCTCC | GTTAACTTTA | CACCGCCAGC | TACTTTGATA | TAAGCATCTT | GTTGTTGTAA | 582 |
| | тасатаатт | 4 Catalah Catalah | CCAAAACAGC | CATCAACAAA | CTTAATCGAT | TATGATCAAT | 588 |

| | | | | | | _ | |
|-----------|--------------|------------|--|------------|------------|------------|------|
| | TATTAAAAGT (| GCTCTGGTTC | CCTCCATGGT | TGCAACAATT | GTTGAACCTG | GAACATTTGT | 6000 |
| _ | TGAACGTTCT 1 | CTAAAAACA | TTTCAGATGG | ATTATTTACA | CCTTTTAATC | CACTTTGCTT | 6060 |
| 5 | CATTTCGAAG A | ATTCcCATTT | CATTCGTTGA | ACCAAAACGG | TTTTTAACAG | CTCGCAAAAT | 6120 |
| | TCGATATGCG T | rggtgttcat | CGCCTTCAAA | ATAAAGCACA | GTATCAACCA | TGTGTTCTAG | 6180 |
| 10 | CAATCTTGGG d | CCAGCAATT | TGACCTTCTT | TCGTTACATG | ACCCACTATA | AAAGTTGCaA | 6240 |
| | TGTTCATTTG T | TTTAGCAATA | TTCATTAAAC | TTTGTGTACT | TTCACGAACT | TGTGAAACAG | 6300 |
| | AACCTGGCGC A | AGAGCTGATT | TCAGGATGAT | ATATTGTTTG | AATCGAATCC | ACTACTAATA | 6360 |
| 15 | AATCAGGTTG T | TCTTCTTTT | ACTGTTTGAT | AAATAACTTC | AAGATCTGTT | TCAGCTAATA | 6420 |
| | CITGCAATIC A | CTTGAATCT | TCATCTAATC | GCTCTGCACG | TAATTTAGTC | TGACTAAGCG | 6480 |
| | ATTCTTCTCC A | GTAATATAT | AGTACTTTTT | TCTTTTGAGA | TAACGATGCA | CAAATTTGTA | 6540 |
| 20 | AAAGTAACGT T | GACTTACCA | ATACCTGGAT | CCCCACCAAT | AAGTACTAAC | GATCCGCTCA | 6600 |
| | CAATACCTCC A | CCTAATACA | CGGTTGAATT | CTGCTGAATC | TGTTAACACT | CTCGGCGTTG | 6660 |
| 25 | TTTCATGTTT A | ATACTATTT | AATTTTTGTA | CTTTACCTGC | TAATTCCTTG | GTTTTAACTC | 6720 |
| 25 | CATGTTTAGG A | TTGGCTGCT | TTTTCAACAA | TTTCCTCCAT | TTGATTCCAA | GCGCCACAAT | 6780 |
| | TAGGACATTT C | CCCATCCAT | TTAGGAGATT | GATAACCACA | AGCCATACAT | TCAAAAATCA | 6840 |
| 30 | CTTTTTTCTT G | GCCArAATT | GCACCTCCAC | TTTCTT | | | 6876 |
| | (2) INFORMAT | ION FOR SE | Q ID NO: 18 | 7: | | | |
| <i>35</i> | (A (B |) LENGTH: | ACTERISTICS 1193 base p cleic acid | | | | |

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

•

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 187:

CAACTCAAAC AGCAGAACAA CGTCGTGAGT TGATTAATGG TGTATTTACT GACATTAATC 60

CCATACATTA AAAATATGAT GTACGTGTTA GCAGATAATA GACATATCTC ATTAATAGCT 120

GACGTATTCA AGGCGTTCCA AAGCTTATAT AACGGACACT ACAATCAAGA TTTTGCAACA 180

ATTGAGTCAA CATATGAATT GAGTCAAGAA GAGTTAGATA AGATTGTCAA ACTAGTAACT 240

CAACAAACGA AGTTATCTAA AGTTATTGTA GATACAAAAA TTAATCCAGA TTTAATTGGT 300

GGATTTAGAG TTAAAGTCGG CACAACTGTA TTAGATGGTA GTGTTAGAAA TGATCTTGTC 360

CAATTACAAA GAAAATTTAG AAGAGTTAAT TAATTATAAA GAGGAGTGAC ATAGATGGCC 420

55

40

45

| ATGTCCGTAA CTGATGT | AGG TACTGTATTA | CAAATTGGTG | ATGGTATTGC | ATTAATTCAC | . 540 |
|--------------------|-----------------|------------|------------|------------|-------|
| GGATTAAATG ACGTTAT | GGC TGGTGAGCTA | GTAGAATTCC | ATAACGGCGT | ACTTGGTTTA | 600 |
| GCCCAAAACC TTGAAGA | GTC AAACGTGGGT | GTGGTTATTT | TAGGACCATA | CACAGGTATT | 660 |
| ACTGAAGGTG ACGAAGT | TAA ACGTACTGGT | CGTATCATGG | AAGTACCAGT | AGGTGAAGAA | 720 |
| CTAATCGGAA GAGTTGT | TAA TCCATTAGGA | CAACCTATTG | ATGGACAAGG | ACCGATTAAC | 780 |
| ACAACTAAAA CACGTCC | agt agagaaaaa | GCTACTGGTG | TAATGGATCg | TAAATCAGTA | 840 |
| GATGAGCCAT TACAAAC | AGG TATCAAAGCA | ATTGATGCTT | TAGTACCAAT | TGGTAGAGGT | 900 |
| CAACGTGAGT TAATCAT | CGG TGACCGTCAA | ACAGGTAAAA | CAACAATTGC | AATTGACACA | 960 |
| ATTTTGAACC AAAAAGA | ATCA AGGTACGATT | TGTATCTATG | TTGCTATTGG | TCAAAAAGAT | 1020 |
| TCAACAGTAA GAGCAAA | NTGT TGAAAAGTTA | AGACAAGCAG | GCGCTTTAGA | CTACACTATT | 1080 |
| GTTGTAGCAG CATCAGG | TTC TGAACCTTCT | CCATTATTAT | ATATTGCACC | ATATTCAGGT | 1140 |
| GTAACAATGG GTGAAGA | AATT CATGTTTAAC | GGTAAACATG | TTTTAATCGT | TTA | 1193 |
| (2) INFORMATION FO | OR SEQ ID NO: 1 | 88: | | | |

25

5

10

15

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5549 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

30

35

40

45

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 188:

60 TGCTAAGAAG TCAAAATAAA CTAACTATNA AACATCTAGT ACGATTATTA AAGTGACAGA 120 AGGTÁCAGGT ATTATTTATG AATAAGTTAA TACTTGGGAT TTATTTATAC CGAATTTTTT 180 CACGAGCATA CTTTTATTTA CCGTTTTTAT TAATTTACTT TTTGATTCAA GGTTATTCCA 240 TAATACAATT AGAAATATTA ATGGCGTCTT ATGGCATTGC AGCATTTTTA TTCTCTCTAT 300 ACAAAGAGAA GTGTTTTAAA ATTTGTAACT TAAAAGATTC TAATAAATTA GTTGTTAGTG 360 420 AAATATTCAA AATCATCGGT TTATTGTTGT TATTATATCA AAATCAATAT TTAATTTTAG TAGTGGCACA AATATTATTA GGGTTAAGTT ACTCAATGAT GGCGGGTGTT GATACCGCAA 480 TAATTAAAAG AAATATAACA AATGAGAAAT ACGTACAAAA TAAGTCAAAT AGCTATATGT 540 TCCTATCATT ATTAATITCA GGGATTATAG GTAGTTATCT TTATGGAATA AATATTAAAT 600 GGCCTATAAT AATGACTGGT ATATTTTCAA TTCTAACAAT TATAATTATT CGATGCACAT 660

| | TACCAGAAGA | GAAGTTTTGG | ATATTGCATT | ATTCTTTTTT | AAGAGCGTTA | ATATTAGGAT | 780 |
|----|------------|------------|------------|------------|------------|------------|------|
| _ | TTTTTATAGG | ATTTATTCCA | ATTAATATAT | ATAATGATTT | AAAACTGAAT | AATTTACAAT | 840 |
| 5 | TTATTTCAGT | ATTAACTTGT | TACACAGTTA | TGGGTTTTGT | ATCTTCACGT | TATTTAACTA | 900 |
| | AATACTTGAA | TTATAAGTTT | GTGTCAGAAA | TTTGTTTAGT | AATATTTTTA | АТААТАТАТА | 960 |
| 10 | CATATCAAAG | TTTCATAGCA | GTTACTATTT | CTATGATATT | TTTAGGTATT | TCTTCAGGGT | 1020 |
| | TAACTCGTCC | ACAAACTATA | AATAAACTTT | CTAGCAGTAG | TAACTTAAGA | GTGATGCTTA | 1080 |
| | ATTATGCAGA | AACGTTATAT | TTTATTTTTA | ATATCGCATT | TTTACTTATG | GGTGGTTACT | 1140 |
| 5 | TATATACAAT | AGGAACTATT | CAATACTTAA | TATTATTTAT | TTCGTTATTA | ATTTTTATAT | 1200 |
| | ATTTAATAAT | AATATTTYAT | TTTACAAGGA | GAGAGCAACA | TGAAAATAAA | AACTGAATTT | 1260 |
| | AAAGGGAACA | ATATACCATA | TGAATACGCA | GCAGGTGCAG | ATGTGAGTGA | TTCTATTAAC | 1320 |
| 20 | GGGAATCCAA | TTAAGTCATT | TCCATTTGAA | GTAATTGAAT | TACCGGAAGG | gactaaatat | 1380 |
| | CTTGCTTGGT | CTTTAATTGA | CTATGATGCA | ATTCCTGTAT | GTGGCTTTGC | TTGGATTCAT | 1440 |
| | TGGAGTGTAG | CTAATGTAAG | TGTTAGTGGC | AATTCAATTT | CTATAAAAGC | AGATTTATCA | 1500 |
| 5 | AGAACAAAGG | GCGACTATGT | ACAAGGTAAA | AATAGCTTTA | CTAGTGGGTT | GTTGGCTGAA | 1560 |
| | GATTTTTCAG | AAATAGAAAA | TCACTATGTA | GGACCTACAC | CACCTGATCA | AGATCATCAA | 1620 |
| o | TATGAATTAA | CAGTTTATGC | GTTAGATCAT | TCTTTAAATT | TGAAGAATGG | GTTCTACTTG | 1680 |
| | AATGAATTTT | TAAAAGAAGT | AAATCAACAT | AAAATTGATC | AAACAAGTAT | TAACCTTATA | 1740 |
| | GGAAGAAAA | TTTAATACTA | AATATCTCAT | CAATATAAAA | TTGTTCAATT | AAAAGTACAA | 1800 |
| 5 | AGAAACAAAG | GTTTTAATTT | ATATATTAGG | TACGGCGTTC | GCTATAATGC | AAAGAAGTAA | 1860 |
| | TTAAATTTAA | GAAATGTAAA | CTTAGTTATT | GTAATGTGAA | TTTATTTGAA | AAAATAGAAA | 1920 |
| | GTATTAACAA | TTATAGCTTT | TACATTAATT | AAAATTTATT | TTTAAAAACA | AGTAAACAAT | 1980 |
| 0 | TTACATACTT | ATAATTTTTG | AAAATTTTÇA | ATTTGTGTTA | TATTGATTTT | GTAAGATACT | 2040 |
| | TTAACTCACA | AAGGAGAGAG | AGTATATGAA | ATTAAAATCA | TTTATAACTG | TAACTTTGGC | 2100 |
| _ | ACTGGGCATG | ATCGCAACGA | CTGGCGCTAC | TGTGGCAGGT | AATGAGGTAT | CTGCAGCAGA | 2160 |
| 5 | AAAGGACAAA | CTACCGGCAA | CTCAAAAAGC | TAAAGAAATG | CAAAATGTTC | CATATACAAT | 2220 |
| | TGCAGTAGAT | GGCATTATGG | CTTTCAATCA | ATCTTACTTA | AATTTACCAA | AAGATAGCCA | 2280 |
| 0 | ATTATCATAT | TTAGATTTAG | GAAATAAAGT | TAAAGCTTTG | TTATATGATG | AACGCGGTGT | 2340 |
| | AACACCTGAG | AAGATTCGAA | ATGCAAAATC | TGCCGTTTAC | ACGATTACTT | GGAAAGATGG | 2400 |
| | ТАСТАЗАЗА | GAAGTGGATC | TTAAGAAAGA | TAGCTACACA | GCAAACTTGT | TTGATTCAAA | 2460 |

| | CAACATGAAG | CATTTAATTT | TACAGTGATG | ATTATAAAAT | AATTGCCTTG | ATACAAAGAT | 2580 |
|----|-------------|------------|------------|------------|------------|------------|------|
| | TACTCGTAAA | TGACATCTTT | GTATTAAGGC | TTTTTCTAAA | TTTAAAAGTG | ATGGGTTAGA | 2640 |
| 5 | GGTCATTGAG | СТТТААААТА | TTCAAAATAC | AAAACATTAA | TGGCCAAAAA | TAAAAGCCGC | 2700 |
| | CTTTATCTGG | GCAGCTTCAA | TAATAAGAAA | GACATATTTC | ATTTTATACT | AAATAGTTAT | 2760 |
| 10 | TGTGATGAAT | CTTTCGGCGG | TTTAATTACT | GCAGCAAAAA | TTGCTGTGAA | AATCGTGAAC | 2820 |
| 10 | AATACTGCCA | TGATAATTGG | ATTCACTACA | TTTAAGCTGT | CTCCACCTAC | TAGGCTATTA | 2880 |
| | AGTACAAAGT | TAACCATTTG | CATTAATAAT | AATGCCCAAA | AGAATGTTAC | GAGGTGTTTC | 2940 |
| 15 | ATGTCATTCT | ACCTCCACTT | TAATTATATA | TATTTTATTT | TAAGTGAAAG | TTAGAAATTT | 3000 |
| | GTATAGTAAC | ATCTCATATA | TTTTGACCAT | ATTATACAGT | TTAAATAAAT | GATTTTATCT | 3060 |
| | GAATGGCTAT | TCTAAATTAA | GCGCATTAAA | ACCAATTTCA | TACTGAAATT | TGACGATAAT | 3120 |
| 20 | AAAGCATTAA | AATTTTATTA | ACTAGTCAAT | ATTCCTACCT | CTGACTTGAG | TTTAAAAAGT | 3180 |
| | AATCTATGTT | AAATTAATAC | CTGGTATTAA | AAATTTTATT | AAGAAGGTGT | TCAACTATGA | 3240 |
| | ACGTGGGTAT | TAAAGGTTTT | GGTGCATATG | CGCCAGAAAA | GATTATTGAC | AATGCCTATT | 3300 |
| 25 | TTGAGCAATT | TTTAGATACA | TCTGATGAAT | GGATTTCTAA | GATGACTGGA | ATTAAAGAAA | 3360 |
| | GACATTGGGC | AGATGATGAT | CAAGATACTT | CAGATTTAGC | ATATGAAGCA | AGTTTAAAAG | 3420 |
| 30 | CAATCGCTGA. | CGCTGGTATT | CAGCCCGAAG | ATATAGATAT | GATAATTGTT | GCCACAGCAa | 3480 |
| 30 | CTGGaGATAT | GCCATTTCCA | ACTGTCGCAA | ATATGTTGCA | AGAACGTTTA | GGGACGGGCA | 3540 |
| | AAGTTGCCTC | TATGGATCAA | CTTGCAGCAT | GTTCTGGATT | TATGTATTCA | ATGATTACAG | 3600 |
| 35 | CTAAACAATA | TGTTCAATCT | GGAGATTATC | ATAACATTTT | AGTTGTCGGT | GCAGATAAAT | 3660 |
| | TATCTAAAAT | AACAGATTTA | ACTGACCGTT | CTACTGCAGT | TCTATTTGGA | GATGGTGCAG | 3720 |
| | GTGCGGTTAT | CATCGGTGAA | GTTTCAGATG | GCAGAGGTAT | TATAAGTTAT | GAAATGGGTT | 3780 |
| 40 | CTGATGGCAC | AGGTGGTAAA | CATTTATATT | TAGATAAAGA | TACTGGTAAA | CTGAAAATGA | 3840 |
| | ATGGTCGAGA | AGTATTTAAA | TTTGCTGTTA | GAATTATGGG | TGATGCATCA | ACACGTGTAG | 3900 |
| | TTGAAAAAGC | GAATTTAACA | TCAGATGATA | TAGATTTATT | TATTCCTCAT | CAAGCTAATA | 396 |
| 45 | TTAGAATTAT | GGAATCAGCT | AGAGAACGCT | TAGGTATTTC | AAAAGACAAA | ATGAGTGTTT | 4020 |
| | CTGTAAATAA | ATATGGAAAT | ACTTCAGCTG | CGTCAATACC | TTTAAGTATC | GATCAAGAAT | 4086 |
| 50 | TAAAAAATGG | TAAAATCAAA | GATGATGATA | CAATTGTTCT | TGTCGGATTC | GGTGGCGGCC | 414 |
| | TAACTTGGGG | CGCAATGACA | ATAAAATGGG | GAAAATAGGA | GGATAACGAA | TGAGTCAAAA | 420 |
| | TAAAAGAGTA | GTTATTACAG | GTATGGGAGC | CCTTTCTCCA | ATCGGTAATG | ATGTCAAAAC | 426 |

| | TGAACCTTAT | AGCGTTCACT | TAGCAGGAGA | ACTTAAAAAC | TTTAATATTG | AAGATCATAT | 4380 |
|----|---------------------|-------------|-------------|------------|------------|------------|------|
| | CGACAAAAA | GAAGCGCGTC | GTATGGATAG | ATTTACTCAA | TATGCAATTG | TAGCAGCTAG | 4440 |
| 5 | AGAGGCTGTT | AAAGATGCGC | AATTAGATAT | CAATGAAAAT | ACTGCAGATC | GAATCGGTGT | 4500 |
| | ATGGATTGGT | TCTGGTATCG | GTGGTATGGA | AACATTTGAA | ATTGCACATA | AACAATTAAT | 4560 |
| 10 | GGATAAAGGC | CCAAGACGTG | TGAGTCCATT | TTTCGTACCA | ATGTTAATTC | CTGATATGGC | 4620 |
| | AACTGGGCAA | GTATCAATTG | ACTTAGGTGC | AAAAGGACCA | AATGGTGCAA | CAGTTACAGC | 4680 |
| | ATGTGCAACA | GGTACAAATT | CAATCGGAGA | AGCATTTAAA | ATTGTGCAAC | GCGGTGATGC | 4740 |
| 15 | AGATGCAATG | ATTACTGGTG | GTACAGAAGC | ACCAATTACT | CATATGGCAA | TTGCTGGTTT | 4800 |
| | CAGTGCAAGT | CGAGCGCTTT | CTACAAATGA | TGACATTGAA | ACAGCATGTC | GTCCATTCCA | 4860 |
| | AGAAGGTAGA | GATGGTTTTG | TTATGGGTGA | AGGTGCTGGT | ATTTTAGTAA | TTGAATCTTT | 4920 |
| 20 | AGAATCAGCA | CAAGCTCGAG | GTGCCAATAT | TTATGCTGAG | ATAGTTGGCT | ATGGTACTAC | 4980 |
| | AGGTGATGCT | TATCATATTA | CAGCGCCAGC | TCCAGAAGGT | GAAGGTGGTT | CTAGAGCAAT | 5040 |
| | GCAAGCAGCT | ATGGATGATG | CTGGTATTGA | ACCTAAAGAT | GTACAATACT | TAAATGCCCA | 5100 |
| 25 | TGGTACAAGT | ACTCCTGTTG | GTGACTTAAA | TGAAGTTAAA | GCTATTAAAA | ATACATTTGG | 5160 |
| | TGAAGCAGCT | AAACACTTAA | AAGTTAGCTC | AACAAAATCA | ATGACTGGTC | ACTTACTTGG | 5220 |
| 30 | TGCAACAGGT | GGAATTGAAG | CAATCTTCTC | AGCGCTTTCA | ATTAAAGACT | CTAAAGTCGC | 5280 |
| | ACCGACAATT | CATGCGGTAA | CACCAGATCC | AGAATGTGAT | TTGGATATTG | TTCCAAATGA | 5340 |
| | AGCGCAAGAC | CTTGATATTA | CTTATGCAAT | GAGTAATAGC | TTAGGATTCG | GTGGACATAA | 5400 |
| 35 | CGCAGTATTA | GTATTCAAGA | AATTTGAAGC | ATAACTATAA | nAATCTTCAG | TAACGTTGTT | 5460 |
| | TTAGTTACTG | AAGATTTTTT | CaGTTTCTTT | ATACTAAGAT | GAGCGACACA | CAATCGTCAT | 5520 |
| | аата а аатат | GAATATTTAT | TAATAATAA | | | | 5549 |
| 40 | (2) INFORMA | TION FOR SE | O TO NO. 19 | 19. | | | |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4832 base pairs
- (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 189:

| AGATTATAGT | AAGATTGATA | GTTTGGCGAC | TGaAGCgCGa | GAAAATTAT | CAGAAGTAAA | 60 |
|------------|------------|------------|------------|------------|------------|-----|
| mCCTTTAAAT | ATTGCACAAG | CTTCTAGAAT | ATCAGGGGTA | AATCCAGCAG | ACATATCTAT | 120 |

55

45

| • | TGGTTAGCAG | AACAATTAAA | AGAACATAAT | ATTCAATTAA | CTGAGACTCA | AAAACAACAG | 240 |
|----|------------|--------------|------------|--------------|-----------------|------------|------|
| | TTTCAAACAT | ATTATCGTTT | ACTTGTTGAA | TGGAATGAAA | AGATGAATTT | GACAAGTATT | 300 |
| 5 | ACAGATGAAC | ACGATGTATA | TTTGAAACAT | TTTTATGATT | CCATTGCACC | TAGTTTTTAT | 360 |
| | TTTGATTTTA | ATCAGCCTAT | AAGTATATGT | GATGTAGGCG | CTGGAGCTGG | TTTTCCAAGT | 420 |
| o | ATTCCGTTAA | AAATAATGTT | TCCGCAGTTA | AAAGTGACGA | TTGTTGATTC | ATTAAATAAG | 480 |
| • | CGTATTCAAT | TTTTAAACCA | TTTAGCGTCA | GAATTACAAT | TACAGGATGT | CAGCTTTATA | 540 |
| | CACGATAGAG | CAGAAACATT | TGGTAAGGGT | GTCTACAGGG | AGTCTTATGA | TGTTGTTACT | 600 |
| 5 | GCAAGAGCAg | TAGCTAGATT | ATCCGTGTTA | AGTGAATTGT | GTTTACCGCT | AGTTAAAAAA | 660 |
| | GGTGGACAGT | TIGTTGCATT | AAAATCTTCA | AAAGGTGAAG | AAGAATTAGA | AGAAGCAAAA | 720 |
| | TTTGCAATTA | GTGTGTTAGG | TGGTAATGTT | ACAGAAACAC | ATACCTTTGA | ATTGCCAGAA | 780 |
| 0 | GATGCTGGAG | AGCGCCAGAT | GTTCATTATT | GATAAAAAA | GACAGACGCC | GAAAAAGTAT | 840 |
| | CCAAGAAAAC | CAGGGACGCC | TAATAAGACT | CCTTTACTTG | AAAAATAATG | CATAATCCTT | 900 |
| | TACAACTAAC | ATAAAAGGAG | CGAATGGATA | ATGAAAAAAC | CTTTTTCAAA | ATTATTTGGT | 960 |
| 5 | TTGAAAAACA | AAGATGACAT | CATTGGACAT | ATTGAAGAAG | ATCGCAATAG | TAATGTTGAA | 1020 |
| | TCCATTCAAA | TTGAACGTAT | CGTTCCCAAC | CGTTATCAAC | CAAGACAGGT | GTTTGAACCA | 1080 |
| 10 | AATAAAATTA | AAGAACTTGC | TGAATCAATA | CATGAACATG | GTTTACTACA | ACCTATTGTT | 1140 |
| | GTAAGACCGA | TTGAAGAAGA | TATGTTTGAA | ATTATTGCTG | GAGAGCGCCG | ATTTAGAGCA | 1200 |
| | ATACAATCAC | TAAATTTACC | TCAAGCAGAC | GTTATTATTC | GTGATATGGA | TGATGAAGAG | 1260 |
| 35 | ACGGCTGTTG | TTGCATTAAT | TGAGAATATT | CAAAGAGAAA | ATTTGTCTGT | TGTTGAAGAA | 1320 |
| | GCGGAAGCCT | ATAAGAAATT | ATTGGAAATT | GGTGATACAA | CGCAAAGTGA | ATTGGCAAAA | 1380 |
| | AGTTTAGGTA | AAAGTCAAAG | CTTTATTGCA | AATAAGTTGC | GTTTATTGAA | GTTGGCGCCG | 1440 |
| 10 | AAAGTACTAC | TTCGCTTAAG | AGAAGGTAAA | ATTACTGAAC | GTCATGCGAG | AgcGGtATTA | 1500 |
| | TCATTGTCTG | ATAGCGAACA | AGAAGCGTTG | ATTGAGCAAG | TCATTGCACA | AAAGCTAAAT | 1560 |
| | GTGAACAGAC | TGAAGATAGA | GTACGCCAAA | AAACGGGGCC | CGAAAAAGTC | AAAGCACAAA | 1620 |
| 15 | ACCTTCGCTT | TGCACAAGAT | GTCACTCAAG | CACGAGATGA | GGTAGGCAAA | AGTATCCAAG | 168 |
| | CGATTCAACA | AACAGGATTA | CATGTTGAGG | : ATAAAGACAA | AGATCATGA | GATTATTATG | 174 |
| | TAAAAATAAA | TCGAATATAT | AAACGTTAGT | r agtaggatgi | CGTATACATC | ATGACTAACA | 180 |
| 50 | CATAAAAGAC | : AAAGCTAAGA | TCATAACAG | TTTGTCTTT | TTTTTTGTT | TACGTGAAAC | 186 |
| | | | | · CTCCTACATE | יייביים ארכים א | CTATGCTCTA | 192 |

| | TTCTAGTCAA | CCTTGCTGGG | GTGGGACGAC | GAAATAAATT | TTGCGAAAAT | ATCATTTCTG | 2040 |
|----|-------------------|-------------------|------------|------------|------------|--------------|------|
| | TCCCACTCCC | TAATTTGAGC | TGGATATACT | TTCATTTGAA | CCCTTTATTG | CTAGTTTATG | 2100 |
| 5 | AAAGTATCAT | GAAAGCTTTA | TGAACATCGC | TTGAGTTGCC | TTTACAGTAG | AAAATTTAAG | 2160 |
| | TTTTACACTT | TGTGTGAATG | ATACGTTTTG | TATTGAATTA | ATTATAGAAA | GGTACGTTGA | 2220 |
| | AGATGTTTTC | AATTGGAAGT | GCAATTCTTC | ATTTTGTCAT | TGGTGGTATC | GCTGTTGCAT | 2280 |
| 10 | TAGCTTCAAT | TATTGCTGAT | AAGGTAGGTG | GTAAGTTAGG | AGGTATTATA | GCTACTATGC | 2340 |
| | CGGCAGTCTT | TCTTGCGGCT | ATTATCGCAT | TAGCTTTAGA | TCATCGTGGT | ACGCAATTAG | 2400 |
| 15 | TGGAGATGTC | GATGAATCTT | AGTACTGGAG | CAATTGTCGG | TATTCTGTCT | TGTATATTAA | 2460 |
| | CIGTATITIT | GACATCTCTC | TACATTAAGC | ATAAAGGTTA | TCGGAAAGGC | GCAATATTCA | 2520 |
| | CAGTTGTTTG | TTGGTTTGTC | ATTTCCCTCG | CAATATTCAG | TATTAGACAT | TTATAGTTTG | 2580 |
| 20 | GAAAATGCGT | GATAATTAGT | TGTATTCAGT | TATTAAGTAA | TAAATTATTG | GAGGCAGAAC | 2640 |
| | ATCATGAAAT | TAACATTAAT | GAAATTTTTT | GTGGGGGGAT | TTGCAGTATT | ATTAAGTTAT | 2700 |
| | ATTGTATCTG | TAACACTACC | TTGGAAAGAA | TTTGGCGGTA | TATTTGCaAC | GTTTCCGGCA | 2760 |
| 25 | GTATTTTTAG | TGTCTATGTT | TATTACAGGT | ATGCAATATG | GTGATAAAGT | CGCTGTGCAT | 2820 |
| | GTAAGTCGTG | GCGCAGTGTT | TGGTATGACA | GGGGTATTAG | TTTGTATTTT | AGTTACATGG | 2880 |
| | ATGATGTTAC | ATATGACGCA | CATGTGGTTG | ATTAGCATTG | TTGTTGGTTT | CCTAAGCTGG | 2940 |
| 30 | TTCATCAGTG | CAGTATĢTAT | TTTTGAAGCG | GTAGAATTTA | TAGCACAAAA | AAGATTAGAA | 3000 |
| | AAGCATAGTT | GGAAAGCTGG | AAAATCGAAT | AGTAAATAGT | GTGAACGTAA | TCTCTTAACT | 3060 |
| 35 | AGGACTAACT | TTGCAAGCAT | TGAATAGCAT | GGAAAAGTTG | CATCATTAAT | AAGTGAAATT | 3120 |
| | CAAGTTGGCA | TTGAGAAAAT | TACAAGCGCG | TAATCATACa | GGTCTGTCTT | AAGGGAGTCT | 3180 |
| | TCGAACCCCG | ATGTTGTCGT | ATGTCAAAAC | ATTTAGTCAA | TCATAAAGGT | GACTTGATTT | 3240 |
| 10 | AACTTTATCT | GATAGTCTGA | TTGTAATGAT | TGTACTAATT | GACTGGAGGC | GTATGTAATT | 3300 |
| - | GAATCTGAGT | AAACAAATTA | AAAAGTATAG | GGAACGAGAT | GGTTATTCAC | AAGAATATCT | 3360 |
| | TGCTGAAAAG | TTATATGTAT | CTAGGCAGAG | TATTTCTAAT | TGGGAAAATG | ACAAAAGCTT | 3420 |
| 15 | ACCAGACATA | CATAACTTAT | TAATGAYGTG | TGAATTGTTC | AATGTAACTT | TAGATGATTT | 3480 |
| | AGTAAAAGGG | ACCATTCCAT | TTGTACCTGA | TATTAAAGCG | CAACGAAGTC | TTAACTTATG | 3540 |
| | GACATATGTG | ATGCTTATTT | TCATGACATT | AGCTGCAATT | TTAATGGGAC | CTTTAGTTGT | 3600 |
| 50 | TTATTGGAAT | TGGACTTGGG | GTGTAACGGT | GGCAATCATT | TTGGGAATAG | GTTTTTATGC | 3660 |
| | АТСТАТСААА | АТАСААСАТТ | TAAAAAAACT | CCATAAAATC | GACAACTACG | ATCC A ATTCT | 3720 |

| | GACAAATGCG | CTTTCTATTA | TATCAGTAAT | TGGTATACTC | AGCCTCATAA | TTTTCCTTAG | 3840 |
|-----|------------|------------|------------|------------|------------|------------|------|
| | TGTGTATTTG | GCAAATAAGT | TTTTATAAAT | CATCGTGGTA | TCGTCTCATA | TTATTTATAT | 3900 |
| 5 | TATCCAAAAT | AGCATAAAAA | AATACCAACA | AGATTTAGAA | CCTTGTTGGT | AATCAAAGCG | 3960 |
| | aTTCATTTAT | AATGAGTCGT | TTTATGTTGT | AAGATTAAAC | AGTTTGTACG | TTAACTGCTT | 4020 |
| 10 | GGTCTCCACG | TTGACCTTCA | GTGATTTCGA | AAGTAACTTT | TTGACCTTCT | TCTAAAGTTT | 4080 |
| ,,, | TGTAGCCATC | GCTAGCGATA | CCTGAGAAAT | GTACGAATAC | GTCTCCGCCA | TTTTCTTGTT | 4140 |
| | CGATGAAACC | AAAACCTTTT | TCTGCTTTAA | ACCATTTWAC | TGTACCGTTA | TTCATATWGA | 4200 |
| 15 | AWACCTCCGT | gTGCTTTTGC | ACTTAATATT | TGTAACAAAT | TCATAACTAA | AAAAGAGGAT | 4260 |
| | ATTCTAAACA | AATACACTAC | AATTTAATTC | ACGAGCTTTT | ATTACGTAAG | ACCAACTATA | 4320 |
| | CGCTCATATT | GGCATAATGT | ACAGTGTTTT | TTGAAAATAA | ATTAAAAAAG | ATTTTTAAAA | 4380 |
| 20 | ACCTTAGAAA | CGTTGATTTA | AAGGGGTTTA | TAAAAATwAw | AAAATTGTAG | TCTTTTATGG | 4440 |
| | TGTTTGCTAG | TTTTCAAAGT | GACATATCGT | TTAAACATGA | TGATTTTATA | AGCAATCCAT | 4500 |
| | AAAAAACAAG | CAGCGATAAA | CGCTACTTGT | TGATATTAAA | ATCTGACTTG | AAAGGTCATA | 4560 |
| 25 | GCAATGTTCT | ATACCGATGG | AATGTGCTTA | CTTGCCTTTT | TCTTCACGAC | GTTTTAAATA | 4620 |
| | ATAAGAGCCA | ССТААТАААС | CAGCTGGAAT | GCCTATCATT | GGTGTTGTGA | ATGAGCTTAA | 4680 |
| | TACAATAACA | AGTATTGTTA | AAGCAATGAC | GTTATACCAA | GTTACAGTCA | AATTTTTCAA | 4740 |
| 30 | ATCCTCATAT | GATTGTTTTA | CTAATTCTCT | AAATTTCATG | ATTCAATCTC | TCCTTTTTTA | 4800 |
| | TAAATCTTTA | GATTGTCAAA | TTAAGCTGGA | CA | | | 4832 |

(2) INFORMATION FOR SEQ ID NO: 190:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5727 base pairs .
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 190:

CAAAGCTGTT CAAAAGGCTT ATAATTTAAA TTTAGATAAC ATACGTACAA TGGAACCTAA 60
GTTGAGATAT CAAGCGATCA ATAAAGGTAA TATTAATTTA ATAGATGCAT ATTCAACTGA 120
CGCTGAATTA AAACAATATG ATATGGTTGT GTTAAAAGAT GATAAGCACG TATTTCCACC 180
ATATCAAGGA GCACCATTAT TTAAAGAAAG CTTTTTAAAG AAACATCCAG AAATTAAGAA 240
ACCGTTAAAC AAACTAGAAA ACAAAATATC TGATGAAGAT ATGCAAATGA TGAACTATAA 300

55

35

40

45

| | GTTAATCAAA | TAACGACCAA | CGCCACATAA | GATGCGTAAC | ACCAAATTAT | ATCTTATGTG | 420 |
|------------|------------|------------|------------|--------------------|------------|------------|------|
| | GCGTTGTTAT | ATTTAAATCT | ATAATTATGT | TCAATTTAAA | CATGCAATAA | TGATTAAAAA | 480 |
| 5 | ATATGACATG | TTAAACACAA | TGTAAGCTAT | TATGATGTGA | AAATAGTAGC | ATTGCATTTT | 540 |
| | AGAAACATAG | AGCGATATAA | TGAATATAAG | TTTTTTGAAA | TTTCAGTTAA | TTCTAAGGAG | 600 |
| | GTTGTTTTTA | TTATGAAAGA | ACAACTTAAT | CAACTATCAG | CATATCAGCC | TGGTTTATCT | 660 |
| 10 | CCAAGGGCAT | TGAAAGAAAA | GTATGGCATT | GAAGGAGATT | TATATAAACT | TGCATCAAAT | 720 |
| | GAAAATTTGT | ATGGACCATC | GCCTAAAGTT | AAAGAAGCGA | TATCAGCACA | CTTAGATGAG | 780 |
| 15 | TTATATTATT | ATCCTGAAAC | AGGATCACCG | ACATTAAAAG | CGGCGATTAG | TAAACATTTA | 840 |
| | AATGTAGATC | AATCACGCAT | TTTATTTGGT | GCGGGATTAG | ATGAAGTTAT | ATTAATGATT | 900 |
| | TCTAGAGCTG | TATTAACGCC | AGGGGATACT | ATTGTTACAA | GTGAAGCGAC | ATTCGGTCAA | 960 |
| 20 | TATTATCACA | ATGCGATTGT | TGAATCAGCT | AATGTGATAC | AAGTACCTTT | AAAAGATGGT | 1020 |
| | GGCTTCGATT | TAGAAGGTAT | TTTAAAAGAA | GTTAATGAAG | ATACGTCATT | GGTATGGTTA | 1080 |
| | TGTAATCCAA | ATAATCCTAC | AGGTACATAT | TTTAATCATG | AGAGCTTAGA | TTCGTTTTTA | 1140 |
| 25 | TCTCAAGTAC | CTCCACATGT | ACCAGTAATT | ATAGATGAAG | CTTATTTTGA | ATTTGTGACA | 1200 |
| | GCAGAGGACT | ACCCGGATAC | ACTTGCTTTG | CAACAAAAAT | ATGACAATGC | TTTCTTATTA | 1260 |
| | CGTACATTTT | CAAAGGCGTA | TGGATTAGCG | GGTTTACGTG | TAGGATATGT | GGTAGCAAGT | 1320 |
| 30 | GAACATGCGA | TTGAAAAATG | GAACATCATT | AGACCACCAT | TTAATGTGAC | ACGTATATCT | 1380 |
| | GAATACGCAG | CAGTTGCAGC | ACTTGAAGAT | CAACAATATT | TAAAAGAGGT | AACACATAAA | 1440 |
| 25 | AATAGTGTTG | AACGCGAAAG | ATTTTATCAA | TTACCTCAAA | GTGAGTATTT | CTTGCCAAGT | 1500 |
| 35 | CAAACGAATT | TTATATTTGT | AAAAACmAAG | CGGGTAAATG | AACTTTATGA | AGCACTTTTA | 1560 |
| | AATGTAGGGT | GTATTACGCG | ACCATTTCCA | ACTGGTGTTA | GAATTACAAT | TGGTTTTAAA | 1620 |
| 40 | GAACAAAATG | ATAAAATGTT | AGAAGTTTTA | TCAAACTTTA | AATACGAATA | GTAAGTGGGG | 1680 |
| - | AGTGGGACAG | AAATGATATT | TTCGCAAAAT | TTATTTCGtC | GTCCCACCCC | AACTTGCATT | 1740 |
| | GTCTGTAGAA | ATTGGGAATC | CAATTTCtCT | TTGTTGGGGC | CCCGCCGGCA | AGGTTGACTA | 1800 |
| 1 5 | GAATTGAAAA | AAGCTTGTTA | CAAGCGCATT | TTCGTTCAGT | CAACTACTGC | CAATATAACT | 1860 |
| • | TTGTAGAGCA | TTGAACATTG | ATTTATGTCT | CAAGCTCAAT | GCAGTGTGAA | TGATGAGGTG | 1920 |
| | AGAGTATTCA | GTGTAAAAAG | CAACAATAGA | TGATATTGTT | TTGTATCAAT | TGCTTTTTTG | 1980 |
| 50 | CTATACTGAA | TCAATACTGA | TATTTTCAGG | agaaga ttaa | AATGACCCGT | AAATCAATCG | 2040 |
| | CGATTGATAT | GGATGAAGTA | TTGGCAGATA | CATTAGGAGA | AATCATTGAT | GCTGTCAATT | 2100 |

| | TTCCTGAACA | TGATGGATTA | ATTACAGAAG | TATTGAGAGA | ACCAGGCTTC | TTCAGACATC | 2220 |
|----|------------|------------|------------|------------|------------|--------------|------|
| | TTAAAGTGAT | GCCGTATGCA | CAAGAAGTTG | TGAAAAAATT | AACTGAACAT | TATGATGTAT | 2280 |
| 5 | ATATTGCTAC | AGCAGCAATG | GATGTACCAA | CATCATTTAG | TGATAAATAT | GAATGGTTAC | 2340 |
| | TAGAGTTCTT | TCCATTTTTA | GATCCTCAGC | ATTTTGTTTT | TTGTGGTAGA | AAAAACATCG | 2400 |
| o | TTAAAGCTGA | TTATTTAATA | GATGACAATC | CTAGACAGCT | TGAAATTTTT | ACTGGTACAC | 2460 |
| | CGATTATGTT | TACAGCAGTG | CATAATATTA | ATGATGATCG | ATTTGAACGC | GTAAATAGCT | 2520 |
| | GGAAAGATGT | AGAACAGTAT | TTTTTAGATA | ATATTGAGAA | АТААААТАТА | TCACTTGAAA | 2580 |
| 5 | AATTTCATGT | AGAAAAGATG | ATGGATAGGC | TATAAAGTAA | TTGTGACTGA | GATGAACTTT | 2640 |
| | TATGTCTTAG | ACACTACAAC | ACTATATTGG | CAGTAGTTGA | CTGCGGGGCC | CCAACATAGA | 2700 |
| | GAAATTGGAT | TCCCAATTTC | TACAGACAAT | GCAAGTTGGG | GTGGsCCCCA | ACATAAAGAA | 2760 |
| 0 | ATACTTTTTC | TTTAGAAATT | AGTATTTCTT | ATGCATGAGT | GTAACTCATG | CATTCATATT | 2820 |
| | TTTAAGTACA | CATTAGCTGT | GACTAATGAT | AAAGAATCGC | TACATAATCA | ATCATTAGTC | 2880 |
| | GTTCTTTATC | ATTTCCGTCC | CGCTCTCAAT | AAATGTTAGT | CTATCTTATT | ATTATAAATC | 2940 |
| 25 | GGATGAATGT | GTTAATCTAT | GGCAGATTAC | ACGTCATCCG | ATTTTTTATA | GAATTTGAAA | 3000 |
| | AAGACGCATA | AACCACTATG | ATTTAAAATA | CAACATCAAT | CATTTTAGTG | gCATGCGCCA | 3060 |
| | AAATTATATG | TCTGTTTTTG | AAACAGGGTA | ATAGCTTAAA | GCTAATAAAA | ACGAATATAA | 3120 |
| 10 | GGTGCGTTGA | ATCTTATGAT | TACACTCCAA | ACCTAATATA | ATATCGGGTT | AAGATCATTC . | 3180 |
| | CGGATGCTTA | CAAATCATTG | ACAGTAAGTA | ACTGAATGGC | ATTTGGTATA | ACCTCAATAT | 3240 |
| 15 | CAATAGGTGT | TTCTAATGAA | ATTTCGCCAT | CAATATCAAC | TTTCATTGCT | GGATCTGTTG | 3300 |
| | TAAGTGAAAT | CTTTTTACCA | GGTATATGCT | CAATACCTTG | AGTAATTTCA | TTCCaATTCA | 3360 |
| | TGCTATCACG | CTTTTTAAAA | ATATCATTTA | AAATACTGAA | ACTTTGTTCA | TTAAAAATGA | 3420 |
| 10 | AAGTGTTCAG | TTCACCATCT | TGAGGAGACA | AATCAGTCAA | TGGTATACGA | CTACCACCAA | 3480 |
| | TGAATGGACC | ATTTGCTGTT | AGTATCATGG | TCGTTTCGCC | AGAATATGTC | TTATCATCTA | 3540 |
| | TTGATAATTG | ATAATTAAAT | TGTGTTGGAT | TTAGCAGTGT | TTTGACAGTT | GATCCAATAT | 3600 |
| 15 | AACTCAATTT | ACCAAATATA | TCTTTTGAAC | CATCTTGTAC | GTTTTCAGCG | TTTTGAACAA | 3660 |
| | TGAGACCTAA | GCCAACAAAG | TTGAGTGCAT | ATTGATTATT | TATTTTAATT | ACATCGTATG | 3720 |
| | TACCAACTIG | TGCAGAAATC | ATTTGTTCAC | TAGCTTGTTT | ATGATTAGGT | GCTATATTTA | 3780 |
| 50 | GCGTTTTTGT | AAAATCATTA | AAAGTACCGC | CTGGTAAAAT | GCCAATAGGG | AGTTGAAGGT | 3840 |
| | CATGTGTCAT | AACACCGTTT | ATAAGTTCGT | TAACCGTGCC | ATCACCGCCA | AGAATAAATA | 3900 |

| | CACCTTCGT | T TTCACTCAA | TGAATAGAA | A GATGCTTAC | A AATTGAACTT | r aatgetgttg | 402 |
|----|------------|----------------|------------|--------------|--------------|---------------------|-------------|
| | TAACTTCCC | C AATACCTTG! | TTAATATTT | TTAATCCAC | T GTGTTCATGO | TAAAAGAGGA | 408 |
| 5 | CACCATGTG | T ATATTTATT | TCCATAGTT | AGCCTACTT | CTAAAAATTO | GTTCATTAAA | 414 |
| | TATATATAC | CACTTTTAA1 | TGTTAATACC | AAAAATATGT | TTTTAAATAG | AGAAAATGGT | 420 |
| 10 | AATAAATGA | A ATTGATTTCT | ATAGAGTGGG | ACGAGAAAA1 | TATAGTTATAG | CTGTCTATAA | 426 |
| 70 | TGAGCATATT | r - AAGTTTTTAT | TTATACTGAT | ' ATCTTGAATT | TAATTAATAG | AAACCTATAA | 432 |
| | AAAAACAGTA | AGCCATTTAA | ATGACTTACT | GTTTTTGAA | TTAGGCCAAC | AATATTAACG | 438 |
| 15 | TATACCTTTC | ATCGCTTTGA | TGATTAAAGG | TGAGAATGCI | ' AATACAATTG | TTGTAACAAT | 444 |
| | AATTGCAACA | ACACCTAGGA | AAATAAAGTA | ATTTGTTTGA | CCTAGTGGTT | CTATTAACTT | 450 |
| | AACTAAAGTA | CCATTGATTG | CTTGTGCAGA | AGCGTTAGTT | AAGTACCAAA | TACTCATCAT | 4560 |
| 20 | TTGGGCATTA | AATGCTTTAG | GTGCTAACTT | AACAGCAGCA | CTATTACCCG | TTGGTGATAA | 4620 |
| | GCATAGCTCA | CCGATAACAC | AAATAATGTA | CGATAAAATA | ACCCAGTTAA | CTGAAAAGTT | 4680 |
| | TGATGAACCT | GATGCATAAC | CTACAATACC | AATTAGTATG | TATGACGCAC | CTGCTAAGAA | 4740 |
| 25 | CGTACCAATT | GCAAATTTTA | CTGGCAGGCT | AGGTTGTTTA | GTTCCAAGCT | TTTGCCATAA | 4800 |
| | AAGTGAAATA | ATTGGAGCTA | GTAATAAAAT | AAATAATGGG | TTAATTGATT | GGAAGATCGC | 4860 |
| | TTCACCAAAG | TTTGTTTTCC | AACCAAATAA | GTTTAATTTC | ATATCTGAAT | GTTCAATTCC | 4920 |
| 30 | ATATATGTTT | AATACATTAG | ACCCTTGTTC | TTGAATAGCC | CAGAACACCA | TTCCAAGAAT | 4980 |
| | AAATAATGGA | ATAAATGCTT | TAACACGAGA | ACGTTCAGTA | TCAGTGACAT | CTTTACTTCT | 5040 |
| 35 | AATAATTAAA | GTGAAGTAAA | TGAnTGGTAA | TGCAATACCT | AATACTAAAA | CAGTATTACT | 5100 |
| | AACTAAGTTA | AATGATAATG | AGTTAGTTAA | TGCACCAATA | ACGATAATTA | ATACAATTGC | 5160 |
| | TAAAACAACA | CTTCCGATAA | TAAGACCATA | CTTTTTCTTT | TCAGCTGGTG | TCAATGGGTT | 5220 |
| 10 | | | | | | ACCATACTAA | 5280 |
| - | ACCTAATGCC | ATACCAACTG | CTGCAATCAA | GAATCCGCCG | TGGAAGTTTT | TAACATTAAC | 5340 |
| | AAAGTGTTGC | AAAATAATAG | GTGATAATAA | TGCACCCATA | TTAACTGACA | TATAGAAAAT | 5400 |
| 15 | AACAAAACCT | GCATCCATAC | GTCTATCATT | TTCAGGATAT | AAACGGCCAA | CGATATTTGA | 5460 |
| | AATGTTTGGC | TTCATTAAAC | CTGAACCAAT | AATGATGAAG | AACATTGATG | TGAATAAGCC | 5520 |
| | GATTAATGCA | AATGGTAAGC | TTAAACAAAT | ATGTCCGATA | ATAATAAAGA | CTGCACCTAA | 5580 |
| 50 | TAAAGTAGCG | CCTCTAGTGC | CTGTAATTCT | GTCAGCAATC | CATCCGCCTG | GTATTGATGT | 5640 |
| | CATATAGATT | AATGAACCAT | Aractgacat | AATTGACATA | CCTCTTCTTT | ፐል ፕሮል አምምርር | 5300 |

(2) INFORMATION FOR SEQ ID NO: 191:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14078 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

10

15

20

25

30

35

40

45

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 191:

TGGACTATTA ACGGCGAAGA AGATTTAACG AAATACTTAC AAACCAATGT TGATGGTATT 60 ATCACAGATG ACCCAGCATT AGCTGATCAG ATTAAAGAAG AAAAGAAAGA CGAAACATAC 120 TTCGATCGTT CTATAGGAAT TTTGTTTGAA TAATATAAAC AAAGACCTCT AAAGTTATCA 180 AGATGATACC TTCAGAGGTC TTTTTAATGT TGCCATCTAT GGGATAGGCA ATCGTTTCAT 240 TCGTTTATAT TCATATGACA AGTATTTGTA TGGCAATTTG GCGTCACAAA CACTTACATG 300 ATTTATTGGT GAATTATTAA TTGTTTTGTG AATGCAAAGG GTTAGAAATT GAATTGTAAA 360 TACTITCTAA TCTTTGTTTC GCTTTAGTCA TTTGATCCAA ATTTTTAGTG CGTATAGCGG 420 ATTTTGCAAT ATAGTGCGCA CTAAAATATC GCGTTTTTGA AACGCATCTA AATTTAGGTA 480 CGATAATTTA TITAAGTCAG TGTTTGCTAT TAATTCATGT AATTGATCTA CAAGCGCTTG 540 ATGTTGATAC GTATGTGATG TAGTTTCAGA TTTGCTTGCT AATTTAATAC CAGTCGTATC 600 AAGGAGCGCC GCTTTAATAC CAGCAACTAA ATATGTTTTG ATTTTCATTT GTGTTGTCAT 660 GCTTTGTTAC TCCTTTGATG TACATTAATC AAAAAAATTA TACACTATTG TATATTGCAA 720 AGCTAATTAA CTATAACAAA AAGATAGTTA ATGCTTTGTT TATTCTAGTT AATATATAGT 780 TAATGTCTTT TAATATTTTG TTTCTTTAAT GTAGATTGGG CAATTACATT TTGGAGGAAT 840 TAAAAAATTA TGAAAAAGCA AATAATTTCG CTAGGCGCAT TAGCAGTTGC ATCTAGCTTA 900 TTTACATGGG ATAACAAAGC AGATGCGATA GTAACAAAGG ATTATAGTGG GAAATCACAA 960 GTTAATGCTG GGAGTAAAAA TGGGACATTA ATAGATAGCA GATATTTAAA TTCAGCTCTA 1020 TATTATTTGG AAGACTATAT AATTTATGCT ATAGGATTAA CTAATAAATA TGAATATGGA 1080 GATAATATTT ATAAAGAAGC TAAAGATAGG TTGTTGGAAA AGGTATTAAG GGAAGATCAA 1140 TATCTTTTGG AGAGAAGAA ATCTCAATAT GAAGATTATA AACAATGGTA TGCAAATTAT 1200 AAAAAAGAAA ATCCTCGTAC AGATTTAAAA ATGGCTAATT TTCATAAATA TAATTTAGAA 1260 GAACTTTCGA TGAAAGAATA CAATGAACTA CAGGATGCAT TAAAGAGAGC ACTGGATGAT 1320 TTTCACAGAG AAGTTAAAGA TATTAAGGAT AAGAATTCAG ACTTGAAAAC TTTTAATGCA 1380

55

| | GTTGTATCA | r attatggtga | TAAGGATTA | r GGGGAGCAC | G CGAAAGAGTT | ACGAGCAAAA | 1500 |
|----|------------|--------------|--------------|-------------|--------------|------------|------|
| | CTGGACTTA | A TCCTTGGAGA | TACAGACAA | CCACATAAA | A TTACAAATGA | ACGTATTAAA | 1560 |
| 5 | AAAGAAATGA | A TTGATGACTT | ' AAATTCAATT | ATTGATGATT | r TCTTTATGGA | AACTAAACAA | 1620 |
| | AATAGACCGA | AATCTATAAC | GAAATATAA1 | CCTACAACAC | ATAACTATAA | AACAAATAGT | 1680 |
| | GATAATAAA | CTAATTTTGA | TAAATTAGTT | GAAGAAACGA | AAAAAGCAGT | TAAAGAAGCA | 1740 |
| 10 | GATGATTCTT | GGAAAAAGAA | AACTGTCAAA | AAATACGGAG | AAACTGAAAC | AAAATCGCCA | 1800 |
| | GTAGTAAAAG | AAGAGAAGAA | AGTTGAAGAA | CCTCAAGCAC | CTAAAGTTGA | TAACCAACAA | 1860 |
| 15 | GAGGTTAAAA | CTACGGCTGG | TAAAGCTGAA | GAAACAACAC | AACCAGTTGC | ACAACCATTA | 1920 |
| 15 | GTTAAAATTC | CACAGGGCAC | AATTACAGGT | GAAATTGTAA | AAGGTCCGGA | ATATCCAACG | 1980 |
| | ATGGAAAATA | AAACGGTACA | AGGTGAAATC | GTTCAAGGTC | CCGATTTTCT | AACAATGGAA | 2040 |
| 20 | CAAAGCGGCC | CATCATTAAG | CAATAATTAT | ACAAACCCAC | CGTTAACGAA | CCCTATTTTA | 2100 |
| | GAAGGTCTTG | AAGGTAGCTC | ATCTAAACTT | GAAATAAAAC | CACAAGGTAC | TGAaTCAACG | 2160 |
| | TTAAAAGGTA | CTCAAGGAGA | ATCAAGTGAT | ATTGAAGTTA | AACCTCAAGC | AACTGAAACA | 2220 |
| 25 | ACAGAAGCTT | CTCAATATGG | TCCGAGACCG | CAATTTAACA | AAACACCTAA | ATATGTTAAA | 2280 |
| | TATAGAGATG | CTGGTACAGG | TATCCGTGAA | TACAACGATG | GAACATTTGG | ATATGAAGCG | 2340 |
| | AGACCAAGAT | TCAATAAGCC | ATCAGAAACA | AATGCATATA | ACGTAACAAC | ACATGCAAAT | 2400 |
| 30 | GGTCAAGTAT | CATACGGAGC | TCGTCCGACA | TACAAGAAGC | CAAGCGAAAC | GAATGCATAC | 2460 |
| | AATGTAACAA | CACATGCAAA | CGGCCAAGTA | TCATACGGAG | CTCGTCCGAC | ACAAAACAAG | 2520 |
| | CCAAGCAAAA | CAAACGCATA | TAACGTAACA | ACACATGGAA | ACGGCCAAGT | ATCATATGGC | 2580 |
| 35 | GCTCGCCCAA | CACAAAACAA | GCCAAGCAAA | ACAAATGCAT | ACAACGTAAC | AACACATGCA | 2640 |
| | AACGGTCAAG | TGTCATACGG | AGCTCGCCCG | ACATACAAGA | AGCCAAGTAA | AACAAATGCA | 2700 |
| 40 | TACAATGTAA | CAACACATGC | AGATGGTACT | GCGACATATG | GGCCTAGAGT | AACAAAATAA | 2760 |
| | GTTTGTAACT | CTATCCAAAG | ACATACAGTC | AATACAAAAC | ATTACGTATC | TTTACAACAG | 2820 |
| | TAATCATGCA | TTCTATGATG | CTTCTAACTG | AATTAAAGCA | TCGAACAATC | GGAAGCATAT | 2880 |
| 45 | TTCTAAATTA | TTTATTCATT | ATAGTCTTAA | ACATAACATG | ACCTAATATA | TTACTAACCT | 2940 |
| | ATTAAAATAA | ACCACGCACA | TCTAAGTGAT | ATACGACAAT | CACAGCAATA | ATAATTGCTT | 3000 |
| | TAGAAAGTCG | TGCCGAACTG | GAACTTACAA | GTCTAGTTCG | AACACACACT | GATGTGAGTG | 3060 |
| 50 | GTTTTCTTTA | TTTTAAACAT | GAACAATCAG | ATAAGTTACT | AGCATTAGCA | AATATTATTA | 3120 |
| | AATCAAAGGG | CTTCGATTCA | TAAAATTTAA | AACAATGATT | AAAATTAGAC | GTGTAAATGT | 3180 |

| | TATTTCACAC | AGCTTCATTA | ATAAAACGAA | ATTGCTTCAA | CCCGCTTCAA | CTTCAACTGG | 3300 |
|----|------------|------------|------------|------------|------------|------------|------|
| | CTTCAACTTC | AGCCTACTTC | ATTCAATAAC | AAAACGAATC | CGCTTCATCC | AAAATCAACC | 3360 |
| 5 | ATTCTAACGC | ACATATTCAA | ATATAGCAGC | TGCACCCATG | CCGACACCAA | TACACATCGT | 3420 |
| | AACCATGCCG | TAACGGCTAT | CGGGACGTCT | ACCCATTTCA | TTAAGTAAAC | GCGCGGTTAA | 3480 |
| | CATTGCGCCT | GTAGCACCTA | ATGGATGACC | TAAAGCAATA | GCGCCACCAT | TCACATTCGT | 3540 |
| 0 | ACGTGATATA | TCTAGACCTA | CTTCTTTAAT | AGATGCAATC | GTTTGAGAAG | CAAATGCTTC | 3600 |
| | GTTCAATTCG | АТСАААТСАА | TGTCTTCAAC | AGATAGATTG | CTGAGTGACA | ATACTTCAGG | 3660 |
| 5 | AATCGCATAT | GCAGGCCCAA | TACCCATAAT | TTTCGGGTCA | ACGCCTACTG | CCTTAAAACC | 3720 |
| 3 | AACGAATCGT | GCAATAGGTG | TCACGCCGAG | TTCTTTCACT | TTATCTCCAG | ACATTAAAAC | 3780 |
| | TACAAATCCT | GCACCATCAG | AAAGTGGGGC | AGATGTTCCT | GCAGTCATAG | TGCCGTCAGC | 3840 |
| 20 | TTTAAATACT | GTACGTAATT | TGGCTAATGC | CTCCATCGTG | GTGTCAGGGC | GTATAAATTC | 3900 |
| | ATCTTGGTCA | AAGATATTTG | TGTGTACTTT | TGGTCCTGCG | TTTGTATATT | CAACTGAGTT | 3960 |
| | TACTTGTATT | GGAATAATTT | CATCTTTGAA | CCGACCATCA | CGTTGTGCGT | CATAGGCACG | 4020 |
| ?5 | TTGATGACTT | CTGACAGCAT | AAGCATCTTG | ATCTTCGCGT | GATACGTCAA | ATTGGGATGC | 4080 |
| | TACATTTTCA | GCAGTTAAAC | CCATAGGATA | TGACGCACCT | ATATCATCAT | ATTGTAAGGT | 4140 |
| | TGGATTGTTT | GTGGGCTCGT | TGCCACCCAT | TGGTACGGCA | CTCATCAATT | CAACGCCACC | 4200 |
| 30 | AGCTACAAGT | ATATCTCCTT | GACCAGCCAT | AATTTGATTG | GCTGCAATCG | CGATGGTTTG | 4260 |
| | TAATCCTGAT | GAGCAGTAGC | GATTCACTGT | TTGACCCGGT | ACCGTGTCAG | ATAATCCCGC | 4320 |
| | ACGCAATGCA | ATCGTTCGTG | CAATGTTTTG | GCCTTGTAAT | CCTTCTGGAA | AAGCCGTACC | 4380 |
| 35 | AACAATGACA | TCTTCAATCA | TATTCTTATT | GAATTTTCCG | TCAATACGTT | TCAATACGCC | 4440 |
| | TTGŢĀATACT | TTGGCTGCGA | CATCATCAGG | TCTTTCGTGG | AATAATGCGC | CTTGCTTTGC | 4500 |
| 40 | TTTCGCTGCG | GCTGAACGCC | CATAAGCTAC | AATGTATGCT | TCTTGCATGG | TTATCATCCT | 4560 |
| | CTCTTAATGA | CTATCTTTTA | ATTACGTAAT | GGCTTACCAG | TTTTTAACAT | ATGTGCAATT | 4620 |
| | CTTTCATATG | ATTTTTAGA | TTTTAGTAAG | TCAATAAAGC | CAATTTTCTC | CAACGATTGA | 4680 |
| 45 | ATGTAACGTT | GATTGATAAA | TGTATTTCTT | GGTAAATCAC | CACCCGCTAA | AATTGTGGCG | 4740 |
| | ATATTTAAGG | CAATATGATA | ATCATGGTCG | CTAATAAAAT | GACCCCGTCT | TTGCGCATCT | 4800 |
| | AATTGTCCTT | GGATCAATGC | TTTGAAGTCT | TCACCTAAAG | CGATATATTG | ATGTCTAGGA | 4860 |
| 50 | TTCGGAATAT | AGTTTGTTTC | TGCTTCATAT | TTCGCACGTT | TGAGCGCAAC | TTCGACACGT | 492 |
| | тстсстатат | TAAAAATAAT | ССТАТСТСТА | TCACGTAAAT | AACCATAACG | ACGTGCCTCA | 4980 |

| | TGTTTGTCAT CAAACTTATG CGATGTGCGT AATATGCGAT CAGCCATTTC TGCAAGC | GCCA 5100 |
|----|---|-----------------|
| | CCGCCACTCG GTAATAAGCC AACACCTGCT TCAACAAGAC CGATATATGT TTCACTT | TGCA 5160 |
| 5 | GCGACAACAA TAGGTGAGTA AAGTACAAGC TCACAGCCAC CGCCTAAGGC ACGACCT | TGA 5220 |
| | ACAGCTGTGA CTACTGGTTT CAAACTATAC TTCAAACGAT TAAAGCTATA ATGTAAT | TTA 5280 |
| | TCAATTGATT GTGCAACGAC ATCATCTACA AGACCGTCTT CATGCGCCTT TTTCATT | 'AAG 5340 |
| 10 | AAAAGGTTAG CACCCACACT GAAATTGTTA CCATCTGCAT AAATAACCAT ACTTGTG | TAA 5400 |
| | TGGTCATTTT CCAGTAAATC AATCGCATCA ACTAACGCAT CGTTGAATTC ATCGGTA | ATG 5460 |
| 16 | ACATTATTTT TACTTTGTAA TTTCAGTAAC AGTTGATCAT CATGAGTTAC GGAAAGT | TTG 5520 |
| 15 | GCATCACCTT TATCCCAAAG TTCATCTTTT ACGAAGTGAG AAATAGGTGT TGCATAT | TCA 5580 |
| | ATGGTCTCAT CTTGTTTATA AAAGCCACCA TCTAAATCAC TAATCCATTG TGGTAAG | TCT 5640 |
| 20 | CCAAGTTCGT CTTCCATACG TGTTTTAACA CGTTCGTATC CCATTGCATC CCATAAT | TGG 5700 |
| | AATGGACCAA GTTTCCAGTT GAACCCCCAG ACAAGCGCAC GGTCTATGTC TCGGAAA | TCA 5760 |
| | TCGGTAGCTT TAGGTACATT GATAGCAGAG TAATAGAAAT TATTACGTAA TGTCTCC | CAT 5820 |
| 25 | AAAAATAGTC CCGCTTCGTC TTGCGCATTG AATATGGTAT CAAGGTTATG CACTAAG | TCT 5880 |
| | TTATTAAATT CATTTAAAAT TGGTAATTGT GGTTGCGATA CAGGTACATA ATCTTGT | FTT 5940 |
| | TCAACATCGT AAACAAGTCG AGCTTTAGTT TCTTTATCCT TTTTGTAAAA TCCTTGT | TTC 6000 |
| 30 | GTTTTACGTC CGAGTGCGCC ATTGTCAAAC AACGTATTTA CAATTTTGAC ATCATGA | AAA 6060 |
| | TAAGGTGTTT CTTCAGGTAC TTGTTGCATG CCTTTAATTA CAGACACTGC AATATCTA | AAA 6120 |
| | CCGACTAGGT CAGATAGCGC ATATGTACCT GTTTTAGGAC GACCAATCGC TTGCCCAC | GTT 6180 |
| 15 | AAAGCATCCA CATCTACAAT GCTTATCTTG TGTTGCTCGG CGCGATACAT AATATCAT | FTC 6240 |
| | ATTCTTTGCG TGCCGACTCT ATTTGCGACA AAGCCAGGCA CATCATTGAC GACAATGA | ACA 6300 |
| | CCTTTACCTA ACACATTTTG CGCGAAATTT TTTACATCTA ATATAATAGA TTCCTTCC | GTG 6360 |
| | TGTGACGTAG GTATTAACTC CACTAATTTC ATAATACGTG GTGGGTTAAA GAAATGTA | AGA 6420 |
| | CCAAAGAATC GTTCTTGATC CTTCTCGTTA AATGCTTGAG CAATCGCATT AATTGGAA | ATA 6480 |
| 5 | CCTGATGTAT TTGTAGCGAA TAAAGCATCT TCTTTAGCAT GTTGTAGAAC TTGTTGCC | CAA 6540 |
| | ACAGCATGCT TAATTTCAAT ATCTTCTTTG ACTGCTTCGA TATATAAATC AGCATCAT | CA 6600 |
| | TTTACCAAGT CATCATCAAA ATTACCATAT GTTAAATGAC TCGCTAGATT TAAGTCGA | AT 6660 |
| 0 | AGTAGCGGCC GTTTCTTATC TGTAATTTTA TCGTAAGATT TTTTCGCAAT GAGATTTG | GA 6720 |
| | TCGTTTTTGT CCACTACAAT ATCTAATAGT TTTACTTTAA GTCCAGCATT CACAAAA | GT 6780 |

| | GTGATTCCTC | CAATTTAGTT | GAGGATAAGA | TAACCATTAA | GATAATTGGA | ATAACGTTGC | 6900 |
|----|------------|-------------|------------|------------|--------------------|------------|------|
| | TATTTTATAA | AATTAATTAA | GTATCTTTGA | CAGTCATCTT | AGCCTCTTAT | TTAAGGAAAA | 6960 |
| 5 | AGCTTTATGC | TTAAAATAAG | TCTTTTTTAG | TGAAATTAAT | GCATCTCATA | TAATTATTTG | 7020 |
| | CTATTTATAC | GAAAGCAGAA | TCTCCAGTCA | AAGCGCGTCC | AATTACTAAG | GCATTAATTT | 7080 |
| | CATGTGTACC | TTCGTACGTG | TAAÀTCGCTT | CTGCATCAGA | GAAGAAACGT | GCAATATCAT | 7140 |
| 10 | AATCGTCAGC | TAGTATGCCA | TTACCACCTG | TAATACCGCG | GCCCATAGCT | ACTGTCTCAC | 7200 |
| | GCAAACGTAA | GGCATTCATC | ATCTTCGCCG | TTGAAGTTGC | AACCTCGTCA | TATTCACCAT | 7260 |
| 15 | GTGCTTGCAT | ATTAGCTAAT | TGAGCACATG | TTGCCATTGC | TTGAGCTAAA | TTACCTTGCA | 7320 |
| | TCATTGCTAG | CTTTTCTTGT | ATTAACTGAT | ATTTACTAAT | TGGTTTGCCG | AATTGCTTAC | 7380 |
| | GCTCAGTGAC | ATAATCTAAT | GTGGCACGTA | AAGCGCCAGC | CATACCACCT | GTAGCCATAT | 7440 |
| 20 | AAGCAACGCC | TGCTCTCGTT | GAATAAAGAA | TTTTGGCAAT | ATCTTTAAAG | CTTGTTATGT | 7500 |
| | TTTGTAAGCG | ATCCGCTTCA | TCTACTTTGA | CATTAGTTAA | TTTAATTAGG | GCGTTAGGAA | 7560 |
| | CAATGCGAAG | TGCGATTTTA | TTATCAATGA | CTTCAATATC | GACGCCATCT | TGTTCTGGTC | 7620 |
| ?5 | TGACTACAAA | GCAATGGGGT | TTGCCAGTTT | CTTTATTTAC | TGCGAATACT | GGAATGACAT | 7680 |
| | CAGATACATG | TGCACCACCA | ATCCATTTCT | TTTCACCATT | GATAACCCAA | GTATCGCCTT | 7740 |
| | GGCGTTCAGC | GACTGTTTCA | AGACCTCCCG | CAACGTCCGA | ACCGTGTTCT | GGTTCAGTTA | 7800 |
| 30 | AAGCAAAGCA | TGTACGCAGT | TCATGTGACT | GTAATTTAGG | TACATATTTC | GCAATTTGTT | 7860 |
| | CTTTGCTACC | TCCGAAATAG | AAAGTGTTAT | GCCCTAAACC | TTGGTGAACA | CCGAGTAGGG | 7920 |
| | TAGCTAAGGA | AATATCAAAT | CGCGCGAGTA | GGTAAGACAT | GAAAAACTGA | AATAGTTGAC | 7980 |
| 35 | TAGGCATTTT | GGCGTTTGGA | CGATCCTTGT | AAAGTAATGG | ATTGTTAAAA | TAATTTAATT | 8040 |
| | CTCCCAGATC | TTTAAAATAG | TCCTCGGGTA | CAGTAGCGTC | TATCCAATGT | TGATTAATAT | 8100 |
| 10 | TTTCACGGTA | CTTACTTTCT | AGCAATGAAT | CTACTTGTTG | TAAAAATTCG | ACTTCACCGT | 8160 |
| | CTGTTAAACC | TTTAGCAATA | CTAAGTACAT | CTTCAGGAAA | TAATGTTTTT | AAGACCGTTT | 8220 |
| | CTTTTTCAAA | TGTCATATAA | ATTCCTCCTA | AAAATAATAT | GAATACTAAT | GTGAAATGCA | 8280 |
| 15 | TTTAATTCAA | AAACAACACG | CTTTATTTGT | AAACGCTTAC | ACTAAATGTC | AAAAATTTTT | 8340 |
| | ATCACCTTTA | AAGTGTTTGC | GAGACTTTGT | CATTCATCAT | TTGTCGAATC | GCAAGTTTAT | 8400 |
| | CTGGTTTCTG | CGTACTGTTT | AACGGCATAT | GTGTCACTGG | TACATACATT | CTTGGGACTT | 8460 |
| 50 | TATAACCTGC | TAAACGACTt | CGCATATGTT | GATTTAAAAT | TTCAGCGTAA | TGAGGTTCAT | 8520 |
| | CTTCCCCAAC | тата атссст | CCACCAATTC | аттеасеата | defendance y dec y | TCATACCCAA | 0500 |

| | AGACATTTTC | GCCACCAGTI | ATGATTAATT | r ctttttttgcg | GTCAATAATA | AATATATCGC | 8700 |
|----|------------|------------|--------------|---------------|------------|------------|-------|
| | CATCGTTGTC | CATCTTCGCT | ' AAGTCACCAC | TTAATAAATA | TCGACCATGA | AATGCTTTGG | 8760 |
| 5 | CAGTCTCTGC | TGGTTTATTC | CAATATCCTC | GCGTGACATT | TTTAGCCTTA | ATTGCAAGTT | 8820 |
| | CGCCAATCTC | ACCAGTAGGT | ACTTCCTCAC | CGTTATCATC | AAGGATACGT | GCATCAACGA | 8880 |
| 10 | ACATGACTGC | TTTACCAATA | CTCATTGGCT | TACGTTTTGA | ATTTTCCGGT | GTATTAACAA | 8940 |
| 70 | GTACAAGAGG | TGCTTCAGTT | AAACCATAGO | CGTTAATAAT | GTTTATGCCA | TATTGTTTAA | 9000 |
| | AAGCTGCTTG | GATACTTGGT | AATGGTTGTG | AACCACCTTG | GATGATATAA | TCCATAGCTC | 9060 |
| 15 | TAAAATTTTC | AGGATTAAAA | TTACTAGCAC | GTAGCGTACT | ATAATACATT | GTCGGAATCA | 9120 |
| | TGATAATAAA | TGTAGGGTGA | TATTGTGCAA | TCATGTCATT | CAATTCTTCG | CCGTTAAAGT | 9180 |
| | AACGTTGAAG | AATAAGTGTG | CCACCTGACA | TTAATACTGG | TAATACAGTA | TCGTTAAACC | 9240 |
| 20 | CTAAAACATG | GAACATTGGT | GTTGATACAA | TCGTAATATA | GTTTGAATTG | AACTTATACG | 9300 |
| | TCAGCTCTAA | GTTTGCACCG | TTATGAACAA | ATGATTCATA | TGAGAACATC | ACACCTTTAG | 9360 |
| | GTGATCCGGT | TGTACCACTT | GTATAAATTA | ATGCTGCAAG | ATCTTGTGGT | TCAACAGGTG | 9420 |
| 25 | TTGCTTGAAA | AGGTTGGTGA | TAATCTGGAT | TTACGATTTC | ATCATATTGC | GCCACATCAA | 9480 |
| | TATCCATATG | CAATAAGTTT | TGGTCAATAT | CGGTGAGTGA | ACTTAAATGT | TTTTCAGCAT | 9540 |
| | AGAAGAGCAG | TTTTAATTGT | GCATCTTCCA | CAATGGCTGC | AATTTCTTTT | GGGTTAAGCC | 9600 |
| 30 | GCCAATTCAA | TGGTAAAAA | ACCGCACCTG | TTTTAAAACA | AGCAAACAAT | AAATCTAATA | 9660 |
| | TTGCAATATC | ATTTGGCGCA | AAAATACCGA | TAACATCGCC | TTTTTTAACA | CCTTGAGATG | 9720 |
| 35 | TTAAATAATG | TGCCATATTA | TCAGCGCGTG | CATTGAGTTG | TTGGTATGTC | CAAGATGTTT | 9780 |
| | GTTTTGCGTG | ATCAATAACG | GCAGGCTTGT | CATCATCGAA | GTCTGAACGC | GTTTTTATCC | 9840 |
| | AATCGAAATT | CATTAGTATA | CCCCCTTTAG | CTTCACTTTC | ATACTTTATG | AATTGATTGT | 9900 |
| 40 | TTAAGTTGTC | CCCATTTTTC | TTTGTAAATG | CTGGTATCAA | TTAATTTTAA | ATGATCAGCA | 9960 |
| | ATAATTGGTT | TAAAAGCCAT | TTGATTCAAA | ATATCTTTAT | GCAAATCAAG | ACCTGGTGCA | 10020 |
| | ATTTCAATTA | GTTTCAAGCC | TTGATTGGTG | AGTTCGAATA | CTGCACGATC | AGTAACAAAA | 10080 |
| 45 | TAGATTTCTT | GCTCGAGTGA | TTGTGAATAT | TGTGCATTAA | AGTCGATATG | GCTCACATCT | 10140 |
| | GATACAAATT | TCTGGTTTTG | TCCTTCAGTT | TCAATGTTTA | ATCGTTGATT | ATGGCATGAG | 10200 |
| | ACATGACTGC | CAGCTACAAA | AGTACCTGAA | AAGATAATTT | TATTTACAGA | TTGCGTAATG | 10260 |
| 50 | TCTATAAAGC | CACCACATCC | ATTTAGTCGG | TCATTGAAGT | AAGACACGTT | GACATTGCCG | 10320 |
| | TATTGATCAA | CCTCAGCAAA | GCTAAGATAG | GCAACTGATA | CACCATTGTT | ATAAATAAAA | 10380 |

| | CGACTCCCAA | CGAATCCACC | GAAAATGCCA | ACATCTAAAA | TCGGTTGCAC | ATCATGTTCA | 10500 |
|------------|------------|------------|------------|------------|------------|------------|-------|
| 5 | ACACATTCTT | CATGCAATAA | ATTAGAGAGT | TCATTATTGA | TGCCATAACC | GATGCTAATT | 10560 |
| | GTATCGCCAT | AAGTTAAAAA | CTGAGCAGCA | CGTCGGAGAA | TCAATTTGCG | ACTATTAAAA | 10620 |
| | GGTAATGCGG | GTTCAGGTAT | TCCATCAATT | CGTTCTTCTC | CAGACAAGGC | TGGTAAATAA | 10680 |
| 10 | TGACTTTGAA | TTACTTGGCG | GTGATTCTTT | TCATCTTCTG | TGACGTATAC | ATAATCGACA | 10740 |
| | AGATTTCCTG | GGATAACAAC | TTCATTCGGT | TTTAGTTGAT | AGTCGTCAAC | TAAAGCTTTA | 10800 |
| | ACTTGTACAA | TAACTTTCCC | ATGATTGGCT | TTCGCGTTTA | ATGCGACATG | ATAACACTCG | 10860 |
| 15 | CTCAAGTACG | CTTCTTGAGT | TAAATAAATG | TTACCTTGTT | GATCTGCGTA | TGTTCCTCTC | 10920 |
| | AGTAGTGCCA | CATCAACGCT | AGGGAATGTG | TAATGTAAGT | ATGTTTCATC | GTTGATGGTT | 10980 |
| | ACTAATGAAA | CTAAATCATC | CGTTGTTCGT | GTATTTACTT | TACCGCCACC | GTATCTAGGA | 11040 |
| 20 | TCAACAGCTG | TGTTTAATCC | GATTTTAGTA | ATAACTCCAG | GTAATAATTG | ATTACTCTGA | 11100 |
| | CGATAATGAG | TTGCAATGAT | ACCTTGTGGT | AAAAAATAAG | CTTCAATGTC | ATTATTTTTC | 11160 |
| | ATTGCTTGTG | CCGTTTTGGA | AGAAGCCGTT | AAAATACTCA | TAATGACACG | TTTAATCATG | 11220 |
| ?5 | CGACGTTCTA | TAAAATCATC | TAAATCCGGT | GCGGCACCTA | AACTATGAAT | ATCATTCGCT | 11280 |
| | AATATAAACG | TTAAATCATT | GGGCGTATGA | TATGTGTCAT | GTTGCGCTAA | CACAGCACGT | 11340 |
| | AGAACTTCGG | CGGGTAAGTT | GGCTACAGCT | AATGCTGGTA | AACCAATCAC | ATCACCATCT | 11400 |
| 30 | TTAATGATAT | GTTGTAAGTC | GTGCCATGTG | ATTTGTTTCA | AGCAAGTCAC | CTCCATCACA | 11460 |
| | TTTGATAAAA | TATAGCGTTT | TTACACTTTG | TGTAAACCCT | Tacaagaaat | ATAACATAAC | 11520 |
| | GACGTTTAAA | ATCAATTAGA | AATATCTTTT | TATTCTGATA | ATAGACACAG | TATAGACACA | 11580 |
| 35 | TTTTGATGGT | CGATAACAAT | TGTAATATCA | AGGGTTTGTA | ATGAATTGAA | TATCATTAAA | 11640 |
| | ATACTTATAT | AAAAATATTG | TTCGGAATAT | AAAAAGTTAA | ATAGGTTTTG | TAAATTTTTA | 11700 |
| 10 | ATGAÄATACA | AAGTGCCCAA | TCGAACAAAG | TATTTATATT | AAAATATGGA | AAATCCATCA | 11760 |
| •0 | ATATTAAATT | AAAATAGTTT | TATTATGAAA | agtgaaagta | GGTAAGTCTA | TGGAAGGTCT | 11820 |
| 1 5 | TAATCATCGA | AGAAATACAG | aaaaagaaga | GACAACACAA | ACGCAATCAG | TTGCACCTAA | 11880 |
| | TACAGGTGAA | GAGGGGATGT | CATCAGCAAG | TACACAATCA | ACTAAGACGT | CCGACATACA | 11940 |
| | TAATGAATCT | ATCGATAAAC | AAATGGAAGC | TAAAGCGCAT | GAAACAGCGC | AAAATACAGA | 12000 |
| | TTTAAAAAAC | GAAGCAAGAA | GTTTATTTGA | TAATGCAACC | AAATCAATCG | GTAGACTAGC | 12060 |
| 5 <i>0</i> | GGGCAATGAT | GAAAGCTTAA | ATCTTAATTT | AAAAGATATG | CTTTCTGAAG | TATTTAAGCC | 12120 |
| | GCATACTAAA | AACGAAGCAG | ATGAAATATT | TATAGCGGGT | ACTGCTAAAA | CTACGCCAGC | 12180 |

| | TTTCACAGTA | ACATTTATTG | GATTATGGGT | CATGGCAGCA | ATTTTTAATA | ACACTAACGC | 12300 |
|----|------------|------------|------------|--------------|------------|------------|-------|
| • | GATTCCGGGT | CTCATTTTTA | TAGGGGCTTT | ' AACAGTACCA | TTATCGGGTT | TGTTCTTCTT | 12360 |
| 5 | TTATGAATCA | AATGCGTTTA | AAAATATTAG | CATTTTTGAA | GTTATTATCA | TGTTCTTTAT | 12420 |
| | TGGCGGCGTA | TTTTCATTAC | TAAGTACGAT | GGTATTATAT | AGATTTGTCG | TTTTTAGTGA | 12480 |
| | TCAATTCGAA | AGGTTTGGTT | CTTTAACATT | TTTCGATGCA | TTTTTAGTAG | GATTAGTTGA | 12540 |
| 10 | AGAAACTGGA | AAAGCACTCA | TTATTGTTTA | TTTCGTCAAT | AAATTGAAAA | CAAATAAGAT | 12600 |
| | TTTGAATGGA | TTATTAATCG | GTGCTGCTAT | TGGTGCAGGG | TTCGCAGTTT | TTGAATCAGC | 12660 |
| 15 | AGGTTATATT | TTGAATTTCG | CTTTAGGAGA | AAATGTCCCA | TTATTAGATA | TTGTCTTCAC | 12720 |
| | ACGTGCGTGG | ACTGCGATTG | GTGGTCATTT | AGTTTGGTCA | kCGATTGTTG | GTGCTGCAAT | 12780 |
| | AGTTATTÇCG | AAAGAACAGC | ATGGCTTTGA | ATTCAAAGAT | ATTTTTGATA | AACGCTTTTT | 12840 |
| 20 | AATATTCTTT | TTATCAGCCG | TTGTTTTACA | TGGCATTTGG | GATACATCTT | TAACTGTACT | 12900 |
| | TGGCAGTGAT | ACGTTGAAAA | TATTTATTTT | AATCGTTATT | GTGTGGATAC | TTGTATTCaT | 12960 |
| | TTTAATGGGG | GCAGGTTTAA | AACAAGTGAA | TTTACTGCAG | AAAGAATTTA | AAGAACAACA | 13020 |
| 25 | GAAAAAAGTA | GACGAATAAT | AATTAAAGCT | TATGTTGCTC | ATATGTTTGT | GACATAAGCT | 13080 |
| | ATTTTTATAA | TTTGTCTTTA | AAAGAGTGGA | ATAGGAATAC | TTTTTGGAGT | TAAAAAAGTG | 13140 |
| | TTtCACGTTA | AACAAATAGT | GACAATTAGA | TTTATATAAA | ATGAACATGA | TTCACTGAAA | 13200 |
| 30 | GTATGTAATA | ATCATTTTAT | TGAAATTCAT | CAAACAGAAA | TTAATACAAT | CATATAAGCA | 13260 |
| | AATTAAACCA | CGCCATAATC | ATATTGGATG | ACTTCGGCGT | GGTTTTTATA | GTTGAAGCAG | 13320 |
| 05 | GGCTGAGACA | TAAATCAATG | TCCCACACTC | CCTTATCGTT | CAATCGTTGT | TCGATAATCG | 13380 |
| 35 | ATTAAATAGA | TACCTTCAGG | TGTTACTTTA | TAATTTTTAA | CCTTAGAGTT | AGCAGCGACT | 13440 |
| | ATTIGATCGT | TGTAAGCAAT | ATAACTGTTT | GGTACATCTC | GACTTGATAA | TTTAATAATA | 13500 |
| 40 | TCATTAGAAA | TATTGTGACG | TTCCTTAACA | TCTACAGTAT | GATTCAATTG | ATTAATTAAA | 13560 |
| | TCATCGACGT | TGCTATTATT | GTAGTCTCCT | TTATTAATAG | CACCATCTTT | TTTATATGCT | 13620 |
| | TGATTAAAGA | AATAACCTGT | ATCTCCACGA | GGAATTGTTC | CGAAACTATA | CATCGTTGCA | 13680 |
| 45 | TCCCATGCAG | AACGGTCTTT | TAAGTAACCT | TCTATGTCAT | CAACACTTTT | AATGTCGATT | 13740 |
| | TCAATATTTG | CTTTTTTAGC | ATCTGATTGT | AATACTTGCG | CAATTTTCGA | TAGCTCTGGA | 13800 |
| | CGACCGTCAT | ACGTAATTAA | CTTAATTTTT | AAAGGGTGTT | CTTTTGTATA | ACCATCTTTA | 13860 |
| 50 | GCTAATAACA | TTTTTGCTTG | TTCGATATTT | TGTTTGGTTA | ACTTAGGTTC | TTTAATATAT | 13920 |
| | GGAATTTTAT | CATTAAATGG | ACTCGTTGCA | GGTTTCGCAT | AACCTTGATA | AATATGATCT | 13980 |

| | TTATLAGTAT GATTATACAT AAGTAAGAAG TTCTAAAn | | | | | |
|----|---|------|--|--|--|--|
| | (2) INFORMATION FOR SEQ ID NO: 192: | | | | | |
| 5 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 486 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | | | | | |
| 10 | (b) Totoboot. Tinear | | | | | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 192: | | | | | |
| 15 | TGAAAACTAA AGTGTTTCTA ATGCGTGACT AAAATTAGTA ATAATTAAGT TCTCATGATA | 60 | | | | |
| | ATAGGTATTT TTGAAAAATG GAGGAGTCTA TAAATGGGTA AAAAAATGGG TCTAGGTTTA | 120 | | | | |
| | TCTATTGCAT TGGTTGTTAT TGGTATTGCC GTTGTATGTT TAATGATTTT TTCTAGTCAA | 180 | | | | |
| 20 | AAAACGACTT ATTTTGGTTA TATGAATAGT AATACAAATG CAGAAAAAGT TGTCAGTGAA | 240 | | | | |
| | AAAGATGGAT TAGTCAAACA TAATATCAAA GTAGAACCAT CTAATGATTT CAAGCCGAAA | 300 | | | | |
| | AAAGGAGACT TTGTAAAATT AGTTTCTAAA GATGATGGGA AGACATTTTA TAAACAAGAG | 360 | | | | |
| 25 | ATTGTTAAAC ATGATGACGT CCCACACGGT TTAATGATGA AAATTCACGA CATGCATATG | 420. | | | | |
| | AATTAATAAA AAAGCATCTA TAACGTAATT TTGAAGAAGT AGAGTTATCT TCTTATGCGT | 480 | | | | |
| | TTTAGA | 486 | | | | |
| 30 | (2) INFORMATION FOR SEQ ID NO: 193: | | | | | |
| 35 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1626 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | | | | | |
| | | | | | | |
| 40 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 193: | | | | | |
| | GAGGTCTATA TACAATTATG GTTGTTCCAG TTAAACGAAC TGATGGCTTT ATTACTAAGT | 60 | | | | |
| | TTAATAGATT AATTGAAAGA CGATTATTAC GTCATTTCAG TAAAAAAGGT TATATCACAT | 120 | | | | |
| 45 | GGGAGGAAAA TTGATTGTCT GACATTTTAA AATGTATCGG TTGTGGTGCG CCACTTCAAT | 180 | | | | |
| | CTGAAGATAA AAATAAACCT GGTTTTGTAC CAGAGCATAA TATGTTTCGT GATGACGTGA | 240 | | | | |
| | TTTGCAGACG TTGTTTCCGC TTGAAAAATT ATAACGAATT CAAGATGTAG GATTAGAAAG | 300 | | | | |
| 50 | TGAAGACTTT TTAAAATTAT TATCAGGACT TGCGGATAAA AAGGGTATTG TCGTCAATGT | 360 | | | | |
| | CGTGGATGTA TTTGACTTTG AAGGATCATT TATTAATGCA GTTAAACGTA TTGTCGGAAA | 420 | | | | |
| | | | | | | |

895

| | TCGAGTTAAA | GAATGGTTAA | AACGAACAGC | AAGAAAATAT | GGTTTGGAAG | CTGACGATGT | 540 |
|----|-------------|------------|-------------|------------|------------|------------|------|
| | CGTATTAATT | TCAGCTGAAA | AAGGCTGGGG | CATAGACGAC | TTATTATCAT | CAATTGCGAA | 600 |
| 5 | TATTCGAGAA | AATGAAGATG | TGTATATTGT | AGGGACAACG | AATGTTGGGA | AATCTACATT | 660 |
| | GATTAATAAA | CTGATTGAAG | CTAGTGTTGG | TGAAAAAGAT | GTAGTAACAA | CTTCAAGATT | 720 |
| | CCCTGGAACA | ACTTTAGATA | TGATAGATAT | TCCTTTAGAT | GAAACATCAT | TTATGTATGA | 780 |
| 10 | TACACCAGGT | ATTATTCAAG | ATCACCAAAT | GACGCATTTA | GTTAGTGAAA | AAGAATTGAA | 840 |
| | AATTATTATG | | | | | | 900 |
| | ATTÁTTCTTC | | | | | | 960 |
| 15 | TTGTTTCTTT | | | | | | 1020 |
| | ATGGCGTAAT | | | | | | 1080 |
| 20 | TAATGAGGTA | | | | • | | 1140 |
| 20 | TGGTCTAGGC | | | | | | |
| | TGTTGAKGTT | | | | | | 1200 |
| 25 | TTATAGGAAA | | | • | | • | 1260 |
| | CTTTAGGATT | | | | | | 1320 |
| | | | | | | | 1380 |
| 30 | ATTAAAGAAA | | | | | | 1440 |
| | GAACGTATCA | | | | | | 1500 |
| | AACACTGTTT | | | | | | 1560 |
| 35 | GTTAAAGGAT | TGCACAGCGT | TTAnCCAGAT | TTAGAAAATG | CATACATTTT | AATTTTGGGC | 1620 |
| | GCAGGT | | | | | | 1626 |
| | (2) INFORMA | | Q ID NO: 19 | 4: | | | |
| | (4) 000 | | | | | | |

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 635 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 194:

AGGGTTAATT GTCGGTTTAA TTGCAATGAA TAAGTTCCAT GTATTAGCTG GCTATAGAGC 60 GAAATTCATC TTAATGGTGA TTTTAACTAT GATGGTCTTC GTACTTATTA ATACGTATTT 120 ACTAAGACAG GTAAAATCTA TCGGTATGTT CTTAATGATT GCTGCATTGG GTCTATACTT 180

55

| GTCTTATATC GAT | AACATGT TCTTCAA | TTA TTTAAATGC | GAGCATCCTA | TAGGCTTGGT | 300 | | | | |
|-------------------------------|-----------------|----------------|--------------|------------|-----|--|--|--|--|
| GCTAGTAATA TTA | ACAGTAC TTGTGAT | TAT TGGCTTTGTA | CTGAACATGT | TTATAAAACA | 360 | | | | |
| CTTTAAGAAA GAG | AGATTAA TCTAATG | TTG ATGAATAGCO | G TGATTGCTTT | AACTTTTTTA | 420 | | | | |
| ACAGCATCTA GCA | ATAATGG CGGACTI | AAT ATTGATGTG | : AACAAGAAGA | GGAAAAGCGA | 480 | | | | |
| ATCAATAATG ATT | TAAATCA ATATGAT | ACA ACGCTATTTA | ATAAAGACAG | CAAAGCGGTT | 540 | | | | |
| AATGATGCGA TTG | CTAAGCA GAAAAAA | GAA CGACAACAAC | : AAATAAAAA | TGATATGTTT | 600 | | | | |
| CAAAATCAAG CGA | GTCACTC GACTCGC | TTG AATGA | | | 635 | | | | |
| (2) INFORMATIO | N FOR SEQ ID NO | : 195: | | | | | | | |
| (i) SEQUENCE CHARACTERISTICS: | | | | | | | | | |
| | | | | | | | | | |

- (A) LENGTH: 13715 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: double(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 195:

| 60 | CATCATTGCG | AAAGATTGTT | ATAAAGTAAT | CTTCTTCATC | TATTATTTGT | CTGAAATGGG |
|-----|------------|------------|------------|------------|------------------|------------|
| 120 | TAGCATACGC | AATTGCCTAG | AACGTATAAT | GCGATCTTGG | TTTTCATTTG | ACGTTGCCAA |
| 180 | ATTTGAATTT | CAATTTCCAC | ACTTGCACCT | CATTCCTTGA | ACATACAAGT | TTTGATTGAA |
| 240 | TCTTGATGCG | CTAAGACAGC | AGCTCACGTG | ATCCAAATTC | TAATAATTTC | TCTTTTATAT |
| 300 | TGCATTCCAA | CTAACGTCTC | TGATTAGTTA | ACATTGAAAA | CCCAGTAATT | CCTCTTTTAC |
| 360 | ATCCCCGAAA | CTACTGCTGG | ACATACATTA | TTGCGCCTTC | TTGGATTTTC | TGAAGTGGTA |
| 420 | AGCATCTAAA | AATCTTCTAC | AAATTAATAT | TTCATGTAAA | CAACTCGTAT | ATAATTCTGT |
| 480 | TTTGTGAGCT | TTTTATCGTT | GCATATACTT | ATTCACTGCC | GAAATAATCC | TGATĀATTAT |
| 540 | TAAACCAAAC | CTGGAAGTCT | TTGAATTCCC | TTGTTGTCTA | CCTCTGTCAT | AATGCGACAG |
| 600 | CATATGGCTT | TAAGTGTTTC | CCTAAATCAA | AGCATCTGCA | CAATTACAAA | TTTTGCTCGC |
| 660 | GAAATCGCTA | TCCTTTAATT | ATAGTCATTC | TAATTCTGTC | GTGTGACAAG | AATGACTTGG |
| 720 | AACCACTCAT | TTGCTTTATT | TATATCCTGG | AAATTCGTTG | TATATTTAAA | ATCCATCGTC |
| 780 | ATCCCAATAT | CGATCTAACA | TTACATTTCT | ACCATTTGTC | AACCTTTTTA | TATAATCTTG |
| 84 | TGCTTTAAAA | GAGACCTTGG | CTGTAATTAC | AAAACATTAT | ACCGTGATAT | CCGATACAAA |
| 90 | ATAAATATCA | TGCTGCATCA | ATTGCGCTTT | AATTTCTTTG | TATTTCAAAA | GTGGTGTATA |
| 960 | тсаатааттс | ACCATTACCT | ATTGCTCTAA | ACATTTTCAA | TATACAC & CATALA | TATCATAAAC |

| | CATTACGCTC | TATCGTTGTG | ACATGAATGT | ' CATCAGATAT | AGAAGCGAAT | TGCATAGAAC | 1086 |
|-----|--------------|------------|------------|--------------|------------|------------|------|
| | TATAGCCGAT | TGCTGTACCA | ATTTCTAAAA | TATTTTTAAC | ATTATTCATA | CGAATTAATT | 1140 |
| 5 | GCTTAATTAA | ATCTAATGTT | AAACGATCTA | CAATTGGCAC | TTCATTTACC | TCGGCAAATT | 1200 |
| | CACGCAAAAC | TTCGATTGAA | CTATTTTGAT | GTTGATGTAA | ATCTATTAAA | TATTTTTTAT | 1260 |
| | TTAGGTCATC | CATGTTTTAA | ACTTCCTTTA | TGTAAAATAA | GTCAATATGA | TTATGACAAT | 1320 |
| 10 | AAAATAAATC | AGCCTTCACA | ATTGATTATA | ATTTTGCCAA | CCAATTAAAT | GACTGATTTC | 1380 |
| | GTGTTAGACG | CAAAGCTATT | TTATTTATAG | AAGCGAATCA | TTCATATAAA | ATTTAACTTT | 1440 |
| 15 | AGATATTTTA | CCATATTTTC | AATAAAATTA | TAAGCGTTAA | TTATTTATAC | ATTGCTTGAC | 1500 |
| | TTAAAAAATA | CTCTTGCCTC | CCCATCTTTA | AGGTTAGCAA | GAGTAAAATC | TTTTTAATTA | 1560 |
| | TTCTTCCATT | TCAGTATTTA | CAACTTCTTC | AATCATGTCC | CATTCTTCAT | CAGTTTCGAT | 1620 |
| 20 | TGGTACTAAC | TTACCACCGT | CACCTGACTC | ATCTGGTTCA | TTGATCATTG | GTACAAGCTC | 1680 |
| | AATCATATCG | TCTTCATCTG | ATTGAGCACC | TTCTTCAGCT | AAGATAACAT | ACTCTTTTTT | 1740 |
| | GAATTCAGGA | TGATAAAATT | CTAAAACTTT | TCGGTATAAA | ACTTCATTTC | CCTCTTCATC | 1800 |
| 25 | GAATAAAGTT | AATAATTCTT | CTTCGTTATT | AATTTCTAGT | TGTGAATCAT | GATTATGTTC | 1860 |
| | AGTCATAGTA | AAATCTCCTT | TTAATGTAGT | GAATCTAAAT | AGCCTTGTAA | AATAAATACC | 1920 |
| | GCTGCCATTT | TATCAATCAC | TTGTTTTCTT | TTTTGTCTTG | AAACATCTGC | TTCTAATAAT | 1980 |
| 30 | GATCGTTCAG | CAGCCATTGT | GCTTAATCTT | TCATCCCACA | TCACAATCTC | AATAGAAGGA | 2040 |
| | TAAGCTTCTA | ATAATTTTTC | TTTATATGTT | AACGAAGCTT | CGCCTCGAAA | TCCTATTGAA | 2100 |
| | TTATTCATGT | TTTTAGGTAG | TCCTATTACG | ACTGTACCCA | CATTATGTTT | TTTAATAATG | 2160 |
| 35 | TCTACTAATT | GGTCAATACC | TAATTCATTA | TTTTCTTCAT | TGATTCGGAG | TGTGTCTAAT | 2220 |
| | CCTTGTGCCG | TCCAACCCAT | TATATCACTA | ATTGCAATTC | CTACCGTTCT | ACTACCGACA | 2280 |
| 10 | TCGAGTCCTA | AAATTTTATG | TTGTAACATA | AATTATTAT | TTTGCTCTTT | TAAATAGTAA | 2340 |
| | GAAACAAGCT- | CTTCCATAAT | AACATCTCTA | TCAATATGAC | GAATTTGATT | TCTTGCTTCA | 2400 |
| | TTTTGGCGTG | GAATATACGC | AGGGTCACCT | GATAATAAAT | AACCTACAAT | TTGGTTTACG | 2460 |
| 15 | GCATTATATC | CTCGTTCATC | TAATGTTCGA | TAAACATTAT | TTAAAACATC | TCTTACATCT | 2520 |
| | TGCGTTGGAA | GTTCTTCATA | GTCGAATTTC | ATTGTTTTAT | CAAAGTTTTC | CATTTGCGAC | 2580 |
| | ACTCCTTTAA | TTACAAATAT | AACTCACTAT | CATCATACAA | TATTATGGCT | TTAAATTATA | 2640 |
| io. | GATTTTTAAT | GTAATCTTTA | ATAAAGCTTA | ATGATTTTGA | GATATTTTCA | GGTTGTGTAC | 2700 |
| | CGCCACCTTG . | AGCCATATCT | GGACGACCGC | CACCTITACC | ACCAACGATT | GGTGCCATTT | 2760 |

| | TCGATACTTT | ATCATCAACA | TTACTTGCAA | GAATGATAAT | TGTATCTTGT | AGTTTAGATT | 2880 |
|----|------------|------------|------------|------------|------------|------------|------|
| | TAAAATCGTC | CATTGTCGAG | CGAATTGCTT | TCGCATTTGG | TACATCCACT | TCAGTAACCA | 2940 |
| 5 | ATACTTTATA | GCCATTGATT | TCTTCAACTT | GATCTTCAAT | ATTACCCATT | TTAAGTGATG | 3000 |
| | TGATTTCTTT | GTCACGTTGC | TCTAATTGTT | TTAATAATGC | TTTTTCTTCA | TCTTGTAATT | 3060 |
| | GTGTTAACTT | ATCGACTACT | TGATCATCAG | ATTTCACTTT | CAGCTGTGAT | TTCATCGTAT | 3120 |
| 10 | TAAATTTCTC | TTGAATATCT | TCTAAATATA | AGAAAGCTGC | TTTACCTGTT | AATGCTTCAA | 3180 |
| | TACGACGCAC | ACCAGCTCCT | GTACCTGACT | CACTTACTAT | TTTGAATAAG | CCAATTTCAG | 3240 |
| | AAGTATTGCG | GACATGAATA | CCACCACATA | ATTCAATTGA | AAATGGTGCC | ATATTTACTA | 3300 |
| 15 | CACGCACAAC | ATCACCATAT | TTTTCACCGA | ATAATGCCAT | TGCGCCCATT | TCTTTAGCTG | 3360 |
| | AAGCAATATC | CATTTCTTGA | ATGTTAACGT | CAATACCTTT | CCAAATITCT | TCATTTACTA | 3420 |
| 20 | AGCGTTCAAC | TTGATCAATT | TCATCATTAG | TCATTGGACC | AAAATGAGAG | AAATCAAAAC | 3480 |
| .• | GTAAACGATC | TGCTTCTACT | AGTGAACCAG | CTTGGTTAAC | ATGATCACCC | AGTACTGATT | 3540 |
| | TCAACGCTGC | ATGTAATAAA | TGTGTTGCAC | TATGGTTCTT | TTGAATGTCA | CGTCGATCAT | 3600 |
| ?5 | TTTGGTTCAC | TTCAGCAGAC | ACTGTAGCGC | CAACATTTAC | TTGGCCAAAT | TGTACTACTC | 3660 |
| | CTTTATGCAA | GTTTTGACCA | TTTGGTGCTT | TGGTTACTTC | ACTAACAGCA | ATTTCAAAAT | 3720 |
| | TGTCATTATA | AACAATACCT | GTATCCGCAA | CTTGTCCACC | ACTGATTGCA | TAAAATGGTG | 3780 |
| 30 | TTTCCGTTAA | CATGAAGTAT | ACTGTTTCAC | CCGCTTCAAC | TTGTGAAACT | TCTTCACCAT | 3840 |
| | TGTATATCAA | GTGTGTTAGT | GTTGTTTGAG | ctGTCGCAGT | ATCATAACCA | ACAAAAGTAC | 3900 |
| | TTGCAGATGT | AATATTTTTC | AATACTTCAC | TTTGAACTTG | CATTGATTGA | GAATTTTGAC | 3960 |
| 35 | GTGCTTGACG | TGCACGATCA | CGTTGTTGTT | GCATTTCTGA | CTCGAATGTT | GTCATATCAA | 4020 |
| | CTTTCAATCC | TGCTTGCACT | GCTATTTCTT | CAGTTAATTC | AATTGGGAAC | CCATACGTAT | 4080 |
| | CATACAATTT | AAATGCATCT | TTCCCATTAA | TTTCATTTGT | TGTCGCTTTA | GCTTTTTTAA | 4140 |
| 10 | TTAATTCATT | TAAAATCGCT | AAACCATCTT | CTAATGTTTC | ATGGAATCGT | TCTTCTTCAG | 4200 |
| | ACTTTATAAC | ACGCTTAATG | AAATCTGCTT | TTTCCTTAAC | ATTTGGATAA | TATGGTTCCA | 4260 |
| 15 | TAATGTCTGC | AACAATATCA | ACAAGTTTGT | ACATAAATGG | CTCATTGATT | CCTAACGTTT | 4320 |
| - | GACTAAAACG | AACGGCACGA | CGTAACAATC | GACGTAATAC | ATACCCTCTA | CCTTCATTGG | 4380 |
| | CAGGTAATGC | ACCATCAGAA | ATTGCAAATG | CAATCGTACG | AATGTGGTCA | GCAATTACTT | 4440 |
| 50 | TAAATGCCAC | ATCTTGTTCG | TTGTTTACTA | AATATTGTTT | ACCTGATACT | TTTTCGATTT | 4500 |
| | CATTCATTAT | AGGCATAAAT | AAATCTGTTT | CATAGTTAGT | ACGTACATTT | TGAGAAACTG | 4560 |

| | TATGATCTTT | ' ATTATGATTG | AATTCACTAA | ATACTAAGTT | CCATACTTCA | AGATAGCGTT | 4680 |
|----|------------|--------------|------------|------------|------------|------------------------|------|
| | CATTTTCTCC | ACCTGGATAC | ATTTCTTCTG | CCGGATCGTC | TTGTCCATAT | GCTTCTCCGC | 4740 |
| 5 | GATCATAGAA | AATCTCAGTG | TTCGGTCCTG | AAGGCCCTTC | ACCAATATCC | CAGAAGTTAC | 4800 |
| | CTTCAATGCG | AATAATACGA | CTTTCTTCAA | GCCCAATATC | TTTATGCCAA | ATGTTGTATG | 4860 |
| | CTTCCATATC | TTCCGGATGA | ATCGTAACGT | ACAATTTATC | TGGCTCCATA | CCCATCCATT | 4920 |
| 10 | TATCACTCGT | TAAAAATTCC | CAAGCAAATT | CAATCGCTTC | TTGTTTAAAA | TAATCACCAA | 4980 |
| | TTGAGAAGTT | ACCTAACATT | TCAAAGAATG | TATGGTGACG | CGCTGTGAAA | CCAACATTTT | 5040 |
| 15 | CAATATCATT | TGTACGAATA | GCTTTTTGAG | AGTTTACAAT | TCTTGGCTTT | TTAGGTGTTt | 5100 |
| | CACGTCCATC | AAAATATTTC | TTTAATGTTG | CTACACCTGA | ATTAATCCAT | AATAATGTAT | 5160 |
| | CATCATCAAT | TGGCACTAAT | GGTGCAGAAG | GTTCAACCAT | ATGTCCTTTT | TCAACAAAGA | 5220 |
| 20 | AATCTAGATA | TTTTTGTCTA | ATTTCACTCG | CTTTTAACTT | TTTCATCATT | TACACATCCT | 5280 |
| | ATTTACTGTT | TTTAAATTAC | CATTCCATAA | AAATTGATGA | CACAGATAGT | CGATTTGCAA | 5340 |
| | AACTAGTATA | AATCAATATC | ATTTTTTATT | ATTAAAAAAT | ÀAAAAACGCC | CATCCTCAAA | 5400 |
| 25 | AGGGACGAAC | GTTATCGCGG | TACCACCCTA | GTTATAAATG | CAATTCAACA | CATTTATCAC | 5460 |
| | TTTAATTCGA | CTATACAGTT | GTGCATAAAG | TAGCGTTCAC | TAATGTTTGT | tGTACTTTTC | 5520 |
| | ACCAACCAGT | ACATCTCTGA | TAAACAAATC | aTTAACTACT | CATCTTTATA | CGAATTTAAT | 5580 |
| 30 | TCTATTTTAG | TTACATTTAC | GCTTGTTGTC | AACGTTCTAT | AAAGTCATAC | GGCGTGATTT | 5640 |
| | CTCCCATATT | AATCATTGGG | TCAATTTTAA | ACATTGTAGC | TTCCGTTAAT | ACATTTGTAT | 5700 |
| 25 | CTGTTTTTGT | TGAATCAGAC | ATAACTTCTT | CACTATCATT | CGATGACATT | GGCGCTTCTA | 5760 |
| 35 | CTTGATCATC | TATTGTCGTT | TGTGAAGCTC | CTGTATCATT | Agttgctgtg | TTTTCCAgCA | 5820 |
| | TTTCTTCATC | TTCTGAATTA | AAATAATTTT | TCAACAATGT | ACATAATTGT | GTTAAACGCG | 5880 |
| 40 | CTTGACCATT | TGTTTTCAAT | CCAATATCAA | ATGCTTCCGG | ATCACCAAGT | AAAACTAAAC | 5940 |
| | TCGTTTTCGC | TCTAGTTAAA | CCAGTATATA | ATATCGGTCT | TTGTAACATT | CTAAAATACT | 6000 |
| | GTTTAACAAT | AGGCATGATA | ACAATAGGAA | ATTCTGAACC | TTGTGATTTA | TGGATTGATG | 6060 |
| 45 | TACAATAAGC | ATGTGTTAAT | TCCATCATAT | CTTGTTTCGT | AAATGTAATT | TCATTACCTT | 6120 |
| | CAAAATCCAC | AACAAGTACA | TCTTTATTAA | GGGCATTTTC | TTTCGCCCAA | AAAATACCAA | 6180 |
| | CAATAACTCC | TATGTCACCA | TTGAATATGT | TATCATTTGG | CCTATTAACA | AGTTGTAATA | 6240 |
| 50 | CTTTGTCACC | TTTTCTAAAG | ACTACATCAC | CAAACTCAAT | TTCTCGTGTG | TCTTTCTTTT | 6300 |
| | TAGGGTTTAA | AATATCTTGT . | AAAACTTGAT | TTAAACGTTT | AATACCGGCA | ىك لا ئىملىدلى بالمليك | 6360 |

| | CTACCTTCTC | AACAACTGTT | GGTATTTGGT | TTGCCTGACA | GTTAATAAAA | CTTCTATCAT | 6480 |
|----|------------|------------|------------|------------|------------|------------|------|
| | GAAAACGCTG | TGTAATATCA | ATTTTCTGAC | CCAACTTCAT | TCGATGTGCT | AATTCTATAA | 6540 |
| 5 | TGCTTGAACC | ATCTTGTTGA | CGATATACTT | CAGTCAGATT | TACTCGTGGT | ATAGCTTTCG | 6600 |
| | ATTCAATTAA | ATCTTTAAAT | ACTTGACCAG | GACCTACAGA | AGGCAATTGG | TCCTCATCAC | 6660 |
| | CTACAAATAT | CAATTGTGCA | TCTAAAGGAA | CTGCACTTAA | AAATTGGTGG | AACAACCAAG | 6720 |
| 10 | TATCTACCAT | AGACATCTCA | TCAATGATTA | TGAGTCGTGC | GTTTATTTCA | TTTTCTAATA | 6780 |
| | TATCCTCTGG | CTTTGTGTCT | TGATTCCAAC | CTATTAAACG | ATGAATCGTC | ATTGCTTCTA | 6840 |
| | ATCCAGTTGA | CTCTTGTAGT | CTCTTAGACG | CTCTTCCTGT | TGGCGCTGCT | AATACAACTG | 6900 |
| 15 | GATAATCATC | ATTGACATAA | TCATCATAAT | CTAATGATAA | GCCATGAATC | TCAGCATATA | 6960 |
| | ATTCAACAAT | ACCTTTAATT | ACTGTCGTTT | TTCCTGTTCC | CGGTCCACCG | GTTAATAGCA | 7020 |
| 20 | TCACCTTAGA | ATTGATAGCC | GTTTGCAAAG | CTTCTTTTTG | TGAAGCTGCA | TAGTTCACTT | 7080 |
| | GATTCGCATC | TTCTATTTCA | CCAATATGCA | TTTGTAAATC | TGACTGTTCA | ATTTCTGTAA | 7140 |
| | GTTTATTTGT | ATGCGTCTTT | ATTCTGAATA | AGTTTTGAAC | ACTITIGATT | TCaGAATAAT | 7200 |
| 25 | ACAAACTTGG | AATTGCAACT | TGTTCaTTGT | CAATAATTAG | TCGTTTTTCC | TCATTTAAGT | 7260 |
| | ATTGCAACAT | TTCGTCTAAT | TTTTCAGGTT | CGATGACCTC | TTCATCTTGa | TAATTTAATA | 7320 |
| | CATCAACCGT | TAAATCTATA | ACAACATTGA | TAGGCAAATA | TGTATGTCCC | TGTTTAATAC | 7380 |
| 30 | ATTCTTCTTC | TAACGTATAG | AGCAACGCAG | CTTTTAATCG | TTCATTATCG | TTATAAGCGA | 7440 |
| | TACCAATATT | TCTAGCAAGT | TGATCTGCTT | TATTAAAACC | AATACCTTTA | ATATCATAAA | 7500 |
| | TCAATTGATA | TGGATTTCGA | TCTAAAATAG | TCAGTGTATC | GCÇGAGATAA | AACTGATAAA | 7560 |
| 35 | TTGCCATTGA | AAGTTTAGGA | CCAAACCCTA | AATCATGTAA | ACGAATCATT | ATTTTTTCAG | 7620 |
| | ATTCTTGATT | TGCTGAAATT | TGTTCTGCAA | TTTGTTTCTG | TTTCTTTTTA | GATAATCCCG | 7680 |
| | AAACTTTTTC | TAGCACTGAA | TGGTCATCTA | ATATATCATT | TATCGCATTG | TCACCTAATG | 7740 |
| 40 | TATTAACAAT | ATTTTGAGCT | GTCTTTTTAC | CTACACCTTT | AAACAAATCA | CTAGATAAAT | 7800 |
| | AACTTATAAT | TGCTTCTTTC | GTTTGTGGCA | TITCITITIC | AAAAGTCTCT | GCTTTTAATT | 7860 |
| 45 | GTTTACCATA | ACGTGGATGA | TCAACAACTT | GCCCTTTAAA | TGTGTAGACA | TCGCCTTCAA | 7920 |
| | CAATATTCGG | AAGAAACCCT | ACAACAGTTG | GCATTGTATC | AAAGTCTTCA | TTTGTTTCAA | 7980 |
| | TAGTATCTAC | TTTAAGCACT | GTATAAAAAT | TATCACTGTT | TTGAAACAAT | ATCGCTTCAA | 8040 |
| 50 | CAGTACCTTT | GATCATTGAA | TAATCAAATA | GTGTAGGGTC | TGACATGTTA | CTCCTCCTCT | 8100 |
| | TTCATTTTAG | TGAATGTTTT | CAGCGCATGC | TGACTTAATA | ACTGTTTAGG | GTCGATAGTC | 8160 |

| | AAGCCCAAAT | TGTATCTTGC | ATCAACATG | A TTTTTATCA | A TCGTTAATAC | ATGTTTAAGT | 8280 |
|----|--------------|--------------|------------|-------------|--------------|------------|------|
| | TGAGTTATGG | CTTCATTAAA | CATTTCTAA | TGACATAAT. | A CAAGACCATA | TTGAAATTGA | 8340 |
| 5 | ACTTCTGCAT | CTTTGTCTTT | ATCTAGTTC | GCAGCAGTC | A TTAAATACGG | CAATGCCAAG | 8400 |
| | CTTAAATGAT | TCTAACTGAT | TAAACGCCAT | ACCGATCAT | A TAATTACAAT | CAACTTGTTC | 8460 |
| | AATCTCTGTT | TGTAATGCTT | GTTGATATAA | TTTAATAGC | TCTTGATAAC | GTTGCTGATT | 8520 |
| 10 | ATAATATACA | TTTGCTAGAT | TAAAAAATAC | GACGCCATTO | TTCGGATCTA | TTGTnAAAGC | 8580 |
| | | | | | | CGATmCCaGC | 8640 |
| 15 | ` | | | | | ATGCTTGTAA | 8700 |
| | CGCTTCTTCT | ATTTTTCCAT | TITGTATGTA | TTGATAAATT | GTTTGTTGAT | CTATCATTTA | 8760 |
| | | | | | | AATCACTTAA | 8820 |
| 20 | | | | | | ATGGAATAGA | 8880 |
| | AATGCTTAAG | AACCATTAAC | GGTTTATTAT | GTAATGGTTC | TTCCACATTA | GCCACCACTA | 8940 |
| | | | | | | GGACAGAAAT | 9000 |
| 25 | GATATTTTAA | CAAAATTAAA | TTCGTTATCC | CCAACTGGCA | TTGCCTGTAG | AATTTCTTTA | 9060 |
| | CGAAATTCTC | TATGTTGTGG | TCCCGCCAAT | ATAACATTGT | AGAGCCTAGG | ACATTGTGAT | 9120 |
| | GTCCCAGACT | CTATCCTCAT | GAATTATTCT | CATCAAAAAC | TGTCTTTCGT | CATTTTCAAC | 9180 |
| 30 | GTTGAAACTT | CAAATAAGTA | ATTTATTGTT | GCCATTGTTT | ATACAACATA | ATTTAATTGA | 9240 |
| | CCTTCATTTT | TGAACACATC | GTCAATTGTT | GCACCACCAA | GACACACATC | ACCTTGATAA | 9300 |
| | AAAACAACTG (| | | | | | 9360 |
| 35 | CAtGGtCGTT ? | ITCACGTTTC | ACAAAAACTT | TCGTATCTTT | TTGGCGATAT | CTAAATTTAG | 9420 |
| | CTGLACATTC A | AAAACCTTGA | TCTAAGTCAT | TATCTTCTGG | ATTTACAAAT | GAATAGTCTG | 9480 |
| 40 | AAGCAATTAA C | GTAATCACTG | TATAATGCAT | CGTGATGGAA | TCCTTGTTCT | ACATATAAAA | 9540 |
| 40 | CATTATCTTT 1 | PAGGTTTTTA | CCGACAACAA | ACCAAGGATC | GCCATCTCCA | CCTATACCTA | 9600 |
| | ATCCATGTCT T | TTGTCCTATT (| GTGTAATACA | TCAAACCACT | ATGTTTACCC | ATTTTCTTAC | 9660 |
| 45 | CATCAAGTGT T | | | | | | 9720 |
| | AGTTTTTTC G | | | | | | 9780 |
| | CTTGTTCTTC A | | | | | | 9840 |
| 50 | CTTTTGAAAG T | | | | | | 9900 |
| | CTACACCACG T | | | | | | 9960 |

| | TTTCTTTATT | ACACATAACG | TCTGGATTTG | GAGTACGACC | TTTTTTGTAT | TCATCTAAGA | 10080 |
|-----|------------|------------|------------|----------------------------|------------|--|-------|
| | AATACGTAAA | GACTTTATCC | CAATATTCTT | TTTCAAAATT | AACAGCGTAA | TACGGAATGC | 10140 |
| 5 | CAATTTGATT | ACACACTTCA | ATAACATCGT | TGTAATCTTC | AGTTGCAGTA | CATACGCCAT | 10200 |
| | TTTCGTCAGT | GTCATCCCAG | TTTTTCATAA | ATATGCCAAT | GACATCATAA | CCTTGTTCTT | 10260 |
| | TTAAGACGTG | GGCTGTTACA | GAACTATCTA | CACCGCCTGA | CATACCAACG | ACAACACGTA | 10320 |
| 10 | TATCTTTATT | TGACAATTAT | GACTCCTCCT | TAAATTTAAA | ATATATTTTA | TGAATTTCAG | 10380 |
| | CTACAATTGC | ATTAATTTCA | TTTTCAGTAG | TCAATTCGTT | AAAACTAAAT | CGAATCGAAT | 10440 |
| . 5 | GATTTGATCG | CTCCTCATCT | TCGAACATTG | CATCTAAAAC | ATGCGACGGT | TGTGTAGAGC | 10500 |
| 15 | CTGCTGTACA | TGCAGATCCA | GACGACACAT | AGATTTGTGC | CATATCCAAC | AATGTTAACA | 10560 |
| | TCGTTTCAAC | TTCAACAAAC | GGAAAATATA | GATTTACAAT | ATGGCCTGTA | GCATCCGTCA | 10620 |
| 20 | TTGAACCATT | TAATTCAAAT | GGAATCGCTC | TTTCTTGTAA | TTTAACTAAA | AATTGTTCTT | 10680 |
| | TTAAATTCAT | TAAATGAATA | TTGTTATCGT | CTCGATTCTT | TTCTGCTAAT | TGTAATGCTT | 10740 |
| | TAGCCATCCC | AACAATTTGC | GCAAGATTTT | CAGTGCCTGC | ACGGCGTTTC | AATTCTTGTT | 10800 |
| ?5 | CACCGCCAAG | TTGAGGATAA | TCTAGTGTAA | CATGGTCTTT | AACTAGTAAT | GCACCGACAC | 10860 |
| | CTTTTGGTCC | GCCAAACTTA | TGAGCAGTAA | TACTCATTGC | GTCGATCTCA | AATTCGTCAA | 10920 |
| | ACTTAACATC | AAGATGTCCA | ATTGCTTGAA | CCGCATCAAC | ATGGAAATAT | GCATTTGTCT | 10980 |
| 30 | CAGCAATAAT | ATCTTGAATA | TCATAAATTT | GTEGCACTGT | GCCAaCTTCA | TTATTTACAA | 11040 |
| | ACATraTAGa | TACTAAAATC | GTCTTATCTG | tAATTGTTTC | TTCAAGTTGA | TCTAAATCAA | 11100 |
| | TAGCACCTGT | ATCATCAACA | TCTAGATATG | TTACATCAAA | ACCTTCTCGC | TCTAATTGTT | 11160 |
| 15 | CAAAAACATG | TAACACAGAA | TGATGTTCAA | TCTTCGATGT | GATAATGTGA | TTACCCAATT | 11220 |
| | GTTCATTTGC | TTTTACTATG | CCTTTAATTG | CCGTATTATT | CGATTCTGTT | GCGCCACTCG | 11280 |
| | TAATATAAT | TTCATGTGTA | TCTGCACCAA | GTAATTGTGC | AATTTGACGT | CTTGACTCAT | 11340 |
| 10 | CTAAATATTT | ACGCGCATCT | CTTCCCTTAG | CATGTATTGA | TGATGGATTA | CCATAATGCG- | 11400 |
| | AATTGTAAAT | CGTCATCATC | GCATCTACTA | CTTCAGGTTT | TACTGGTGTG | GTCGCAGCAT | 11460 |
| 15 | AATCTGCATA | AATTTCCATG | TTTGGACACT | CCTCACAATT | TTATCAATGT | TCCAATAATA | 11520 |
| | GCACCTTACA | TACTATTTT | CTACTTTTCT | GTTTAACTTT | ATTTATAATG | TTTTTAATTA | 11580 |
| | TATTTTACCA | TTTTCTACAC | ATGCTTTTCG | ATAGGCTTTT | TTAAGTTTAT | CGCTTTATTC | 11640 |
| 50 | TIGICITITI | TATAAATTTT | AGTATTTGCA | GATATTTTTT | TATTTGTAAA | ATGTAACGTA | 11700 |
| | CTATTATTTT | CCTTATCACC | ТАТТААТТАА | עייים איירידעבעידע איירידע | ATTCCATTCC | The transfer of the same of th | 11760 |

| | GTCCCTATTC | GAGAAGGTGA | AGATGAACAA | ACAGCAATTA | ATAATATGGT | TAATCTCGCA | 11880 |
|----|------------|------------|------------|------------|-------------|------------|-------|
| | CAACATTTAG | ACGAATTATC | ATATGAAAGA | TATTGGATTG | CTGAACACCA | TAACGCTCCC | 11940 |
| 5 | AACCTAGTAA | GTTCAGCAAC | TGCTTTATTA | ATTCAACATA | CGTTAGAACA | TACGAAACAC | 12000 |
| | ATACGTGTAG | GTTCTGGAGG | CATCATGTTA | CCTAATCATG | CTCCATTAAT | CGTTGCGGAA | 12060 |
| | CAATTTGGCA | CGATGGCAAC | ATTATTTCCA | AATCGTGTCG | ATTTAGGATT | AGGACGTGCA | 12120 |
| 10 | CCTGGAACAG | ATATGATGAC | CGCAAGTGCA | TTAAGACGAG | ATCAACATGA | TGGTGTTTAT | 12180 |
| | AAATTTCCAG | AAGAGGTTTC | ATTATTACAA | CAATATTTCG | GCCCTGCTCA | CCAACAAGCA | 12240 |
| 15 | TATGTTCGTG | CTTATCCAGC | AGTAGGTAAA | AATGTGCCTT | TATACATTCT | TGGTTCTTCA | 12300 |
| 15 | ACAGATTCTG | CACATTTAGC | TGCTCGCAAA | GGGCTTCCAT | ATGTGTTCGC | TGGACATTTT | 12360 |
| | GCACCTCAAC | AAATGAAAGA | AGCTATCGAA | ATTTACAAAA | CGTTATTTGA | ACCTTCTGAT | 12420 |
| 20 | GTATTAGACG | AACCTTATGT | TATTGTATGT | TTAAATACAA | TCGTTGCTGA | AAATGATGAC | 12480 |
| | GAAGCACAAT | ATTTAGCTTC | ATCTATGGCA | CAAGTAATGG | TTAGTATCAC | TCGTGGCAGA | 12540 |
| | ATGCAGCCCG | TTCAACCGCC | AACACATGAA | CTACAAAATA | TATTAACGCC | GAGAGAATAC | 12600 |
| 25 | GCGATGGCTA | TGGAAAGACA | GAAAATATCA | TTAATAGGTT | CAGAAAATAC | TGTTCAACAA | 12660 |
| | AAAATTCAAG | ATTTTATGGA | AACTTATGGT | GAAGTCAACG | AAATTATGGC | AATAAGTTAT | 12720 |
| | ATTTATGATA | AAGATATGCA | ATTAGACTCT | TATCGTCGGT | TCAAGAATGT | TATAAATCAG | 12780 |
| 30 | ATAAATGAAA | AAAACACTTT | ATAATGTGAT | AAATAAACTA | agtgaaagta | TGTATCCATA | 12840 |
| | ATATTAATAA | AAATATACAG | TAACAGCATT | TTGAATGAAA | GATGTCTTTA | TTGTTCAATC | 12900 |
| | ATTTATTTTA | GTAATGATTC | AAATTCACTT | AAAATYCTAA | tGCAAATATG | AAAGCGCCCC | 12960 |
| 35 | TTCAcTTTAC | ACTGTGTAAG | TGTTTATTTG | ATGGGGCGCT | TTCAAAATAT | TGAAAAGCAT | 13020 |
| | ATCCAAAATT | TAAAGAAATT | TATTTCTCTT | TATCTTCATT | TTCTTTTTTC | TCTTCGTTAT | 13080 |
| 10 | TCGATCCTGT | ATATTCATTT | ATCTTATCTT | TTACATTTTT | AACTTGTTCA: | TTATCGCTAT | 13140 |
| | TTTTAAATTT | TTCTACGCGT | CTTTAGCTTT | ATCCATAAAA | CTCATATTAA | TCGCTCCTCT | 13200 |
| | TATATTTGAT | TAGTTTAATT | GAACTTATTT | TTTAAGTTTA | TCAATTGCAT | CAGTTATTTT | 13260 |
| 15 | GTTTTTAGCA | TTTTCAACAA | CTTCTTTTGC | TTTaCCAGTC | GCTTTATCTT | GCTGACCTTC | 13320 |
| | TTTTTCTAAT | TCTTTGTTAT | CAGTAACGTT | ACCTACTGTT | TCTTTAACAT | TTCCTTTAAA | 13380 |
| | TTGATCGAAC | TEACTTTCGT | CTGCCATAGT | GAAACCTCCT | TGGATGTATA | TATTTATATA | 13440 |
| 50 | CCACTAAGGA | GGTTCGCTmm | mCAyymyAAT | ATGAAGTTTT | TATGTTATAG | TATAGTATTT | 13500 |
| | ATACGATTAA | ATATAAAACA | TGTATCCGTC | TAAATCTTCA | CTTGTATCTA | CATATTCCGC | 13560 |

| | TAGTIGITIT TGCGCAGGTG GTTCTGATTC AATACTTTCA ACAAATGTAA TTGGACCTTC | 13680 |
|----|---|-------|
| | TAACAGTCTT ATAATATCCC CTGCTGAGAT TTCTT | 13715 |
| 5 | (2) INFORMATION FOR SEQ ID NO: 196: | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 873 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 15 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 196: | |
| | AAATCCATAA TGTCATGATA ATCTGCATAT GCTTCATATA ATTCAATCAT TGTGAATTCA | 60 |
| | GGGTTATGTC TAGTTGATAC ACCTTCATTA CGGAATACTC TACCAATTTC ATATACTTTT | 120 |
| 20 | TCAAGTCCAC CGACAATTAM ACGTTTTAAA TGCAACYCAA TAGCAATACG CATGTATAAC | 180 |
| | GTTGCATCTA ATGCATTATG ATGTGTTACA AATGGTCTAG CAGCTGCTCC ACCAGCAATT | 240 |
| | TGGTGCATCA TAGGTGTTTC TACTTCCAAG AAACCTTTAT TATTTAAATA ATTACGCATT | 300 |
| 25 | TCTTGAATGA TTTTACTACG ATTAATAAAT GTACGAGTGC TATCTTCGTT CGTAATTAAA | 360 |
| | TCTAAATATC TTTGACGATA LCTCTGTTCA ATATCCTGTA AACCGTGGAA TTTATCCGGT | 420 |
| | AATGGTCGCA ATGATTTAGT TAGTAGCGTG AATTTCTTCG CTTTAACCGA TAATTCGCCA | 480 |
| 30 | GTATTTGTTT TGAACATTAC ACCTTCAACA CCAACGATAT CGCCTAAATC AGCATTTTTC | 540 |
| | CATAAATCAA ATTCGTCATC GCCAACTTGA TCTTTACGAA CGTAAATTTG AATTTGTCCA | 600 |
| | GCTAAGTCCT GAACGTGTGC AAATCCTGCT TTACCTTTAC CACGCTTAGT CATTAATCGT | 660 |
| 35 | CCAGCTATAG CGACATGACT ATCCGCTTCT TTTTCTACCA ATTCTTCTTT AGAATACTGG | 720 |
| | TCCEACTCTT CTTTCAAATC ACTAGATAAA CCTGAACGGT CAAATTTAGA ACCAAACGGG | 780 |
| 40 | TCTATACCAA GATCATATAA TTCTTGTAAT TETTGACGTC GAACCAACAT TTGGTCATTC | 840 |
| | ATTTCTTCTG ACATAACTET CTCTCCTTTA ACT | 873 |
| | (2) INFORMATION FOR SEQ ID NO: 197: | |
| 45 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 452 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 50 | | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 197: | |

| ACCATAATAT | GAATGGCTTC | AGGATCAAAA | TAAAGACCAA | CTTCACTGCC | TACTTCAGCT | 120 | | |
|-------------------------------------|------------|------------|------------|------------|------------|-----|--|--|
| TTTTTAGTCG | TTTGTATTAC | CCATTCATAA | CCTTTATTGT | CTATACAACA | TATTTCATAG | 180 | | |
| TGGACCCCTC | TAAATAACAT | AGAATCAACA | GTTGCTTTAA | ATAATCCTTC | TTCAGCTTTG | 240 | | |
| ATTAATGATA | TATCTTCTGG | TCGAATAACG | ACTTCTACTT | TTTTATTTTC | AGGAATACCC | 300 | | |
| ATATCGACAC | ATTCGAAATC | TTGCCCATAA | ATATTCACGA | CATAATCTCT | AACCATGCGC | 360 | | |
| CCTTCAACAA | TATTAGATTC | TCCAATAAAA | TCAGCTACAA | ATCGATTCAC | TGGTTCGTCA | 420 | | |
| Tatatatctg | TTGGTGTGCC | AAATTGTTGA | AT | | | 452 | | |
| (2) INFORMATION FOR SEQ ID NO: 198: | | | | | | | | |
| | | | | | | | | |

15

10

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2308 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double (D) TOPOLOGY: linear

20

25

30

35

40

45

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 198:

| TF | AGGTTGGGT | TCTAACATAC | GATAAAGCTC | AACAAATCAA | CACAGCTTTC | TTTGTAAAAT | 60 |
|----|------------------|------------|------------|------------|------------|------------|-----|
| TC | STITAATAC | TGCATTAGCA | GAACGTGATT | ATTATTTAA | TATAGATGGA | ACAAATGCTT | 120 |
| T | PAGATTATT | TAATGCTGAA | GGTGATGGTG | TTGGGGGATT | AACAATCGAC | AATTACGATG | 180 |
| GI | CATTTGTT | GATTCAATGG | TACTCAAAAG | GTATTTATAA | ATTTAAATAT | GCCATTCTTG | 240 |
| A | AGCGGTTAG | AAAAGTATTT | GATTATAAAT | CTATTTACGA | AAAAGTAAGA | TTTAAAGACA | 300 |
| G | GAATATAG | TGGTGGTTTT | GTTGAAGGAG | ATGCACCTGa | GTTTCCAATT | GTTATCGAAG | 360 |
| Ą | AACTTCAC | ATTTTATAAT | GTAGACCTTG | AAGATGGTTT | GATGACAGGT | ATCTTTTTAG | 420 |
| A? | CAAAAAGA | AGTGCGCAAG | AaATTAAGGG | ATCAATATGC | CAAAGAACGC | CATGTTTTAA | 480 |
| A | CTTATTTAG | TTATACAGGT | GCTTTTTCTG | CAATAGCAGC | AAGTGAGGCA | TCTTCAACAA | 540 |
| C | AAGTGTAGA | TTTGGCTAAT | CGTTCTCGTA | GTTTAACTGA | AGAAAATTTT | GGATTAAATG | 600 |
| C. | PATTGATCC | TAAATCCCAA | TATATTTATG | TCATGGACAC | TTTTGATTTC | TATAAATATG | 660 |
| C. | rgcacgaca | TGGACATAGT | TATGACACGA | TCGTGATTGA | TCCACCTAGC | TTTGCGCGTA | 720 |
| A | CAAAAAACG | TACATTTTCA | GTGCAAAAAG | ATTATGACAA | ATTAATTAAT | GGCGCCTTAA | 780 |
| A: | TATCTTATC | ATCTGAAGGA | ACATTATTGT | TATGTACAAA | CGCAAGTGTA | TATCCATTAA | 840 |
| A | GCAATTTAA | AAATACTATT | AAAAAGACGC | TTGAAGAGAG | TGGCGTTGAT | TATGAATTAA | 900 |
| C. | TGAAGTTAT | GGGATTACCA | AAAGATTTTA | AAACGCATCC | ACATTATAAG | CCATCTAAAT | 960 |

| | TATTGAGAAA | AAGAAGGGTG | ATAATATTAT | GGGATTCAAA | AACAATTTAA | CATCAAATTT | 1080 |
|-----|------------|------------|------------|------------|------------|------------|------|
| | AACAAATAAA | ATCGGTAATT | CAGTCTTTAA | AATAGAAAAT | GTTGACGGAA | AAGGTGCAAT | 1140 |
| 5 | GCCAACGACG | ATTCAAGAAT | TGAGAGAAAG | ACGACAACGT | GCTGAAGCAA | TTGTAAAGAG | 1200 |
| | AAAGTCTTTA | ATGTCATCAA | CAATGAGCGT | TGTTCCAATT | CCGGGTTTAG | ATTTTGGTGT | 1260 |
| | TGATTTAAAA | TTAATGAAAG | ATATTATCGA | AGATGTTAAT | AAAATTTATG | GTTTAGATCA | 1320 |
| 10 | TAAGCAAGTT | AATAGCCTTG | GGGATGATGT | GAAAGAAAGA | ATTATGTCTG | CAGCAGCAAT | 1380 |
| | TCAAGGTAGT | CAATTTATTG | GTAAAAGAAT | TTCAAATGCA | TTTTTAAAAA | TTGTAATTAG | 1440 |
| 15 | AGATGTAGCT | AAACGTACTG | CTGCAAAACa | AACAAAATGG | TTTCCTGTTG | TAGGACAAGC | 1500 |
| | TGTGTCTGCA | TCTATTAGTT | ACTATTTTAT | GAATAAAATT | GGAAAAGATC | ACATTCAAAA | 1560 |
| | ATGCGAAAAT | GTTATTAAAA | ATGTCATGTA | GGTGCTATAA | TAGTTTTGCA | ATTTGCAAAT | 1620 |
| 20 | TTTACTGAAA | CCGGTTTTAA | ACGAATTGAA | TTTAAAGcAT | GGTTTTGGTA | AAGTTAATGT | 1680 |
| | АТААААСТАА | GTTAGyATTG | TAATAATATK | GAAGATTCTA | ACTATACGAA | GGAGAAATGT | 1740 |
| | AATTATGGAA | CAAAATTCAT | ATGTAATCAT | CGACGAGAmT | GGTATTCACG | CTAGACCAGC | 1800 |
| 25 | AACAATGTTA | GTACAAACAG | CTTCAAAATT | CGATTCTGAT | ATTCAATTAG | AATATAACGG | 1860 |
| | TAAGAAAGTA | AACTTAAAAT | CAATCATGGG | TGTTATGAGC | CTTGGTGTTG | GTAAAGATGC | 1920 |
| | TGAAATTACA | ATTTATGCTG | ACGGTAGTGA | TGAATCTGAC | GCCATTCAAG | CAATCAGTGA | 1980 |
| 30 | CGTCTTATCA | AAAGAAGGAT | TGACTAAATA | ATCATGTCTA | AATTAATTAA | AGGTATTGCC | 2040 |
| | GCATCTGATG | GTGTCGCAAT | TGCTAAAGCT | TATTTATTAG | TTGAGCCAGA | CTTAACATTC | 2100 |
| | GACAAAAATG | AAAAAGTCAC | TGATGTTGAA | GGAGAAGTTG | CAAAGTTCAA | TAGCGCTATC | 2160 |
| 35 | GAAGCTTCTA | AAGTTGAGTT | AACTAAAATT | AGAAATAATG | CAGAGGTTCA | ACTAGGTGCT | 2220 |
| | GATAÃAGCTG | CTATCTTTGA | TGCAcaTTGG | GGGGTGGTAG | ATGACCCTGA | ATTAATTCAA | 2280 |
| 40 | CCAATCCAAG | ATAAGATTAA | AAATGAAA | | | | 2308 |
| . = | | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 199:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5559 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 199:

AAGTAATAAA TCGTCTCATT TGGCAACTGA CGCATAATTT CTTTAGCTAC TGTCAAACCT 60

55

45

| | TTTAATTTTT | AGTTTATCAT | AACTAAGCAT | TGGATTTTAG | TATTATGCAC | TGTGTTTACC | 180 |
|----|------------|------------|------------|------------|------------|------------|------|
| | ATTTTTGTCA | TTATAATATT | TATTTTAAAT | CAGCCCACTA | TCATATTGTC | ATGTAATCTG | 240 |
| 5 | CTTATTAAAA | AAATCCCTTC | CAAGTTATTG | TGTATCTCCA | TTCAATTTAA | TTTTGAAAGG | 300 |
| | AACATAACWT | TTTAACTCAA | AAGGGATTAA | TTTnTAnTCT | ACTTCATGGT | CTGAACCAAA | 360 |
| | GAATGATTTA | AACATGTGGA | ATGTTGTTTC | TCTGTTCATT | GCTGCAATGG | ATGTTGTTAA | 420 |
| | TGGAATACCT | TTAGGGCAAG | CATTAACACA | GTTTTGTGAA | TTACCACACT | GCTGTAAGCC | 480 |
| | ACCAGTACCC | ATTAATGCAT | TTAAACGTTC | ATCTTTAGTC | ATAGATCCTG | TTGGGTGCAA | 540 |
| 15 | ATTAAACAAA | CGAACTTGCG | AGATTGCTTG | TGCACCAACG | Aatttattat | TTTCAGTAAC | 600 |
| 5 | ATTAGGACAA | ACCTCTAAAC | ATACACCACA | TGTCATACAT | TTAGATAATT | CATAAGCTGT | 660 |
| | TTGACGTTTT | TTCTCTGGCA | TACGTGGTCC | CGGACCTAAA | TCATACGTTC | CATCAATTGG | 720 |
| 20 | GATCCATGCT | TTCATACGTT | TTAAGTTATC | GAACATTCTA | GAACGATCAA | CTTGTAAGTC | 780 |
| | ACGGATAACT | GGGAAAGTAT | TCATTGGCTC | TAAACGAATA | GGTTGTTCTA | ATTGATCAAC | 840 |
| | AATCGCAGAA | CAAGATTGTC | TTGCACGACC | ATTGATAACC | ATAGAACATG | CTCCACATAC | 900 |
| 25 | TTCTTCTAAG | CAGTTCATAT | CCCAGACAAC | AGGTGTTGTT | TTTTCACCTT | TAATATTAAC | 960 |
| | TGGGTTACGT | CTAATTTCCA | TTAAACAAGC | AATGACGTTT | AAATTTTCAC | GATATGGAAT | 1020 |
| | TTCAAATGTT | TCTTCATAAG | GCTTAGAATC | ACTTGTATCT | TGTCGTTTAA | TAATTAATTT | 1080 |
| 30 | TACTGTTTTT | TGTTTCGGTT | TAGATTGTGT | TTCATGTTGT | GGAGTGTTTT | TCACTGATTG | 1140 |
| | TTCAGTCATT | ATTTTTTACC | CCCTTTAGAC | TTACTTGTGT | AATCACGTTT | ACGAGGTGGT | 1200 |
| | ATTAAACTCA | CATCGACGTC | ATCATAAGTA | AACTGCGGTT | TTTCAAATGC | GCCTTGGAAT | 1260 |
| 35 | GAGGCCATTG | TCGTTTTTAA | CCACTCTTCA | TCATTACGCT | CTGGGAATTC | TGGTTTATAA | 1320 |
| | TGGGCACCGC | GTGATTCGTT | ACGGTTATAT | GCACCAATCG | TAATAACACG | TGCAAGTACT | 1380 |
| 10 | AACATGTTCC | ATAGTTGACG | GGTAAAGAAT | ACCGCTTGGT | TACTCCAAGT | TTGAGTATCT | 1440 |
| 10 | TCCATATCAA | TATCTTCATA | ACGTTTCATC | AATTCAACAA | TCTTTTTATC | TGTTTCTAAC | 1500 |
| | AGTTTTTCAT | TTTCACGAAC | AACAGTTACA | TTTGCTGTCA | TAATTTCACC | AAGTTCACGG | 1560 |
| 15 | TGTAATTTAT | ATGCATTTTC | TGTACCGCGC | ATAGCTAATA | ATTTATCAAA | ACGTTCTTGC | 1620 |
| | TCTTCAGCTT | TACGCTTTTC | AAAAATACTT | TCGTCCATAT | CAGTATATGA | TCGATCAATA | 1680 |
| | TTTGAAATAT | AATCAATCGC | GTTTGGACCT | GCTACTGTAC | CACCATAAAT | CGCTGATAAC | 1740 |
| 50 | AATGAATTGG | CACCTAAGCG | GTTACCACCA | TGTTGAGAGA | AGTCACATTC | TCCAGCTGCA | 1800 |
| | AATAACCCTT | таататттст | CATTTGATCA | ТААТСТАСАТ | ATAGACCACC | CATTGAATAG | 1860 |

| | TAAATCTCAA | TGATACCACC | TAGTTTTACA | TCTAACTCAT | GTGGATCTTT | ATGTGACAAA | 1980 |
|----|------------|------------|--------------------|---------------|---------------|-------------|------|
| | TCAAGATATA | CCATGTTTTC | GCCATTTATA | CCTAATTTTT | GGTTAATACA | TACATCGAAA | 2040 |
| 5 | ATTTCACGCG | TTGCGATATC | ACGAGGTACT | AAGTTACCAT | AATCAGGATA | TTTCTCTTCT | 2100 |
| | AAGAAGTACC | AAGGCTTACC | ATCTTTATAT | GTCCAAATTC | GTCCACCTTC | ACCACGTGCT | 2160 |
| | GATTCACTCA | TTAGTCGCAG | TTTATCATCA | CCAGGGATTG | CAGTAGGATG | AATTTGAATG | 2220 |
| o | AACTCACCAT | TAGCATAAAT | AGCGCCTTGT | TGGTAAACAA | TGGAAGCCGC | TGATCCTGTA | 2280 |
| | TTAATCATTG | AGTTTGTTGT | TTTACCGAAA | ATAATACCAG | GGCCACCCGT | TGCCATAATA | 2340 |
| _ | ACTGCATCTG | AACCAAATGT | TTCAATCTCA | GCAGTTGTCA | TATTTTGTGC | AnCGATACCT | 2400 |
| 5 | CTTGCACTAT | CATCGTCACC | TTTAACTATG | CCAAGGAATT | CCCATCCTTC | ATACTTCGTA | 2460 |
| | ACTAATCCAT | CTACTTCATA | TGCACGAACT | TGTTCATCCA | ATGCATATAA | TAATTGTTGT | 2520 |
| 0 | CCAGTTGTTG | CCCCTGCATA | TGCTGTTCTG | TGATGTAATG | TACCACCGAA | ACGTCTAAAA | 2580 |
| | TCTAATAGAC | CTTCATTTGT | TCTATTGAAC | ATTACGCCCA | TACGGTCTAA | TAAATGAATA | 2640 |
| | ATTTTAGGTG | CTGCCTCTGT | CATCGCTTTA | ACAGGTGGTT | GGTTTGCAAG | GAAATCGCCA. | 2700 |
| 25 | CCATACACTG | TATCATCAAA | GTGAATCCAA | GGAGAATCGC | CTTCCCCTTT | AGTATTGACC | 2760 |
| | GCACCATTAA | TGCCACCTTG | GGCACAAACA | GAGTGCGAAC | GCTTTACTGG | TACAACTGAG | 2820 |
| | AACAAATCTA | CATGTGCACC | TTTTTCTGCC | GCTTTAATTG | TTGACATTAA | GCCCGCTAGG | 2880 |
| 80 | CCACCTCCGA | CAACAATAAG | ATGTTTCTCT | GCCATAAAAA | TTTCACTCCC | CTAAATTTTC | 2940 |
| | AATCTATATT | TGTTAAATGC | GATGTATTAC | ATAAAGGCAA | TAATTGCAGT | AACACCAATA | 3000 |
| | TACGAAATAA | CTAAAAATAC | GATTAATGAA | ACCCATGTAA | ATACTCGTTG | TGATTTTGGA | 3060 |
| 35 | GATTGAAGTC | CACCCCAAGT | AACTAAGAAT | GACCATAAGC | CATTTGCAAA | GTGGAACACA | 3120 |
| | ACAGCAATAA | TACAAATAAT | ATAAAATATT | GCCCATCCAG | GATGTTGCAA | TGTTTCGTGC | 3180 |
| | ATTÄAATCGT | AATTCACTTC | TTTGCCGTAA | AATGCTTTTT | GTAAACGTGT | TTGCCATAAA | 3240 |
| 10 | TGGATACCAA | TAAAGATAAA | TGTTAAGATA | CCACTCACTC | TTTGGAAGAA | GAACATCCAG | 3300 |
| | TTTCTAAAAA | TCGAGTAATG | TCCAACATTT | TCTTTTGCTG | TAAATGCAAT | GTGTATACCA | 3360 |
| 45 | AACAAACCGT | GATATAACAA | CGGAATGTAT | ATAAATAAA | ATTCTACAAT | AATTAGAAAT | 3420 |
| | GGTAATGATT | CCATAAAGTT | AGATGCCTTA | TTAAACGCTT | CAGCACCTTG | TGTTGCTTGG | 3480 |
| | TGATTCACTA | ATAAATGAAC | GACCAAAAAT | GCACCTATTG | GGATAATACC | TAATAACGAG | 3540 |
| 50 | TGAATACGTC | TTAGATAAAA | TTCATTTTTT | GATTGAGCCA | AAAGGAGTCC | CCCCTGTGAA | 3600 |
| | | | OF A THOUSAND A TO | THE REPORT OF | CTT 3 8 CCCCC | тааастсата | 3660 |

| | CGATCACCAA | ACTGCATGTC | GAACAATGTA | ACATTTGGAT | TCGATATTTA | AAATTGCTTG | 3780 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | TGATGATAAA | CTTTCTCATT | TAGAAAACGC | TTCCACGTAC | ATTCAAAAAA | ATAACTTTGT | 3840 |
| 5 | TAACCATATT | GTAACATTAT | TTCATATATT | TTGGGGCATG | AGAATGATTC | TCACGCCCAG | 3900 |
| | TAATTTATTT | ATGCAATTGT | TCATGTAGGT | TCTTTGCGAC | GTTTTCAGGA | ATACCTATAT | 3,960 |
| | TTTTAAAATC | TTCAAGTGTA | GCTTCCTTCA | TTTTCTTGAT | TGAACCGAAT | GAACGCAATA | 4020 |
| 10 | ATAATGTTTT | ACGTTTGTTA | CCGATACCAT | CTATATCATC | AAGTATTGAT | TTCAAGCCTG | 4080 |
| | TCTTTTGACG | TGTTTGTCTA | TGAAATGTGA | TTGCGAATCT | GTGAACCTCA | TCTTGGATAC | 4140 |
| | GGTGCAACAA | ATAAAATGCC | TGGCTATTTT | TCTTCAGTGG | TACAATTTCT | GCACTAGCGC | 4200 |
| 15 | CATATAATAA | TTCAGATGTT | TGGTGTTTAT | CATTTTTCTG | CAAACCTGCA | ACAGGGATAT | 4260 |
| | CAAGACCTAA | TTCGTTTTGT | AGCACATCAA | TAACCCCGTT | CATATGTCCT | TTACCACCAT | 4320 |
| 20 | CTACTATTAT | TAAATCAGGT | AATGGTAATC | CTTCGTTTAA | AACGCGAGAA | TATCGTCGTC | 4380 |
| | TTACTACTTC | TCTCATTGAT | TTGTAATCAT | CTGGACCTTT | AACCGTTTTG | ATTTTATACT | 4440 |
| | TTCTATAATT | TTTCTTATCT | GGTTTACCGT | CGACAAATGT | AACCATTGCT | GACACTGGAT | 4500 |
| 25 | CCACACCTTG | AATATTAGAA | TTATCGAATG | CTTCAATTCT | AATTGGTGTT | TGAATTCCCA | 4560 |
| | TTTGTGTTCC | AAGTTCTTCA | ATAGCTTTAA | TCGTTCTGGA | CTCATCACGT | GATATTAATT | 4620 |
| | CAAATTTATT | ATTTAAGGAT | ACTTTAGCGT | TATGTGCAGC | TAGGTCAACC | ATATCTTTTT | 4680 |
| 30 | TGGGACCTCG | CGCGGGTTGA | ACGATTTTAG | TGTCCACAAC | AGATTGAATC | ATTTCTTTAT | 4740 |
| | CCAAATTACG | TGGTACATGA | ACTTCCTTAG | GTAAAATATG | TTGGTTTAAG | CTATAAAATT | 4800 |
| | GTCCAATAAA | TGTATAAAAT | TCTTCTTcTT | CTGTTTGCTG | TAATGGAATC | ATCGTTGTAT | 4860 |
| 35 | CTCGCTTTAT | CATATTACCT | TGTCGTATAA | AGAAAACTTG | GATACACATC | CATCCTTTAT | 4920 |
| | CAACACTATA | ACCAAAGACA | TCACGAATCG | TTTTATCTGA | TGACATAATT | TTTTGTTTGT | 4980 |
| 40 | TTGTCAGATT | TTGAATATGT | TGAATTAAAT | CTCTATATTC | TTTAGCCCGT | TCAAAATCAA | 5040 |
| 40 | GTGATTCACT | TGCAGTTAAC | ATTCGCTCTT | CTAAACTTTT | TAAAATTGTT | TTGTCTTCCC | 5100 |
| | CATTCAGAAA | ATCAGTAATT | TCCTTCGTCA | TTTGTGCGTA | TTTACTCAAA | TCAACGTCAT | 5160 |
| 45 | ATACACATGG | TCCTAAACAT | TGTCCAATAT | GGTAATAAAG | ACATAATTTA | TCTGGCATCT | 5220 |
| | TATCACATTT | GCGATATGGA | TATATTCTGT | CTAATAACTT | TTTAGTTTCT | TGAGCAGAAT | 5280 |
| | ATGCATTCGG | ATACGGTCCG | AAATATTTGC | CAGTACCTTG | TTTTACAGTT | CTCGTCACTA | 5340 |
| 50 | GTAGTCTAGG | ATATTTCTCC | TTCGTAATTT | TAATAAATGG | ATAACTTTTA | TCATCCTTTA | 5400 |
| | ATAATATATT | ATATCTTGGT | TGATATTGTT | TAATCAGATT | CAATTCCAGT | AAAAGTGATT | 5460 |

| | TTTTAGCATC ATGAGCACCC GTAAAATATG ATCGCAATC | 5559 |
|-----|--|------|
| | (2) INFORMATION FOR SEQ ID NO: 200: | |
| 5 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 4594 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 10 | | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 200: | |
| 15 | AAATCAATCG AGTGGCATGT CAAGGTCATA TCAATATTTT AGAATCTGCG ACTATGAGAG | 60 |
| ,,, | AGGAAATAAA TGAAATTGCG CGACGTATCA TCGTTGATAT TCGTGATAAG CAATTACGAT | 120 |
| | ATCAAGATAT TGCTATTTTA TATCGTGATG AATCTTATGC TTATTTATTT GATTCCATAT | 180 |
| 20 | TACCGCTTTA TAATATTCCT TATAATATTG ATACAAAGCG TTCGATGACA CATCATCCGG | 240 |
| | TCATGGAAAT GATTCGTTCA TTGATTGAAG TTATTCAATC TAATTGGCAA GTGAATCCAA | 300 |
| | TGCTACGCTT ATTGAAGACT GATGTGTTAA CGGCATCATA TCTAAAAAGT GCATACTTAG | 360 |
| 25 | TTGATTTACT TGAAAATTTT GTACTTGAAC GTGGTATATA CGGTAAACGT TGGTTAGATG | 420 |
| | ATGAGCTATT TAATGTCGAA CATTTTAGCA AAATGGGGCG TAAAGCGCAT AAACTGACCG | 480 |
| | AAGATGAACG TAACACATTT GAACAAGTCG TTAAGTTAAA GAAAGATGTC ATTGATAAAA | 540 |
| 30 | TTTTACATTT TGAAAAGCAA ATGTCACAAG CGGAAACTGT AAAAGATTTT GCAACTGCTT | 600 |
| | TTTATGAAAG TATGGAATAT TTCGAACTGC CAAATCAATT GATGACAGAG CGAGATGAAC | 660 |
| | TTGATTTAAA TGGTAATCAT GAAAAGGCGG AGGAAATTGA TCAAATATGG AATGGCTTAA | 720 |
| 35 | TTCAAATCCT TGATGACTTA GTTCTAGTAT TTGGAGATGA ACCAATGTCG ATGGAACGTT | 780 |
| | TCTTAGAAGT ATTTGATATT GGTTTAGAAC AATTAGAATT TGTTATGATT CCGCAAACAT | 84 |
| | TGGACCAAGT AAGTATTGGT ACGATGGATT TGGCTAAAGT CGATAATAAG CAACATGTTT | 90 |
| 40 | ACTTAGTAGG TATGAATGAT GGAACGATGC CACAACCAGT AMTGCGTCAA GCTTGATTAC | 96 |
| | AGATGAAGAA AAGAAATACT TIGAACAGCA GGCTAATGTC GAGTTAAGTC CAACATCAGA | 102 |
| | TATTTTACAG ATGGATGAAG CATTTGTTTG TTATGTTGCT ATGACTAGAG CTAAGGGAGA | 108 |
| 45 | TGTTACATTT TCTTACAGTC TAATGGGATC AAGTGGTGAT GATAAGGAGA TCAGCCCATT | 114 |
| | IGITACHITI ICITACAGIC INATOGNIC MATAGIANI GITTINGGIC. IGIGGETTI | |

TITAAATCAA ATTCAATCAT TGTTCAACCA ATTGGAAATT ACTAACATTC CTCAATACCA

TGAAGTTAAC CCATTGTCAC TAATGCAACA TGCTAAGCAA ACCAAAATTA CATTATTTGA

AGCATTGCGT GCTTGGTTAT ATGATGAAAT TGTGGCTGAT AGTTGGTTAG ATGCTTATCA

1200

1260

1320

55

| | GTTTGACAAT | GAAACTGTAA | AATTAGGTGA | AACGTTGTCT | AAAGATTTAT | ATGGTAAGGA | 1440 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | AATCAATGCC | AGTGTATCCC | GTTTTGAAGG | TTATCAACAA | TGCCCATTTA | AACACTATGC | 1500 |
| 5 | GTCACATGGT | CTGAAACTAA | ATGAGCGAAC | GAAGTATGAA | CTTCAAAACT | TTGATTTAGG | 1560 |
| | TGATATTTTC | CATTCTGTTT | TAAAATATAT | ATCTGAACGT | ATTAATGGCG | ATTTTAAACA | 1620 |
| | ATTAGACCTG | AAAAAAATAA | GACAATTAAC | GAATGAAGCA | TTGGAAGAAA | TTTTACCTAA | 1680 |
| 10 | AGTTCAGTTT | AATTTATTAA | ATTCTTCAGC | TTACTATCGT | TATTTATCAA | GACGCATTGG | 1740 |
| | CGCTATTGTA | GAAACAACAC | TAAGCGCATT | AAAATATCAA | GGCACGTATT | CAAAGTTTAT | 1800 |
| 15 | GCCAAAACAT | TTTGAGACAA | GTTTTAGAAG | GAAACCAAGA | ACAAATGACG | AATTAATTGC | 1860 |
| | ACAAACATTA | ACGACAACTC | AAGGTATTCC | AATTAATATT | AGAGGGCAAA | TTGACCGTAT | 1920 |
| | CGATACGTAT | ACAAAGAATG | ATACAAGTTT | TGTTAATATC | ATTGACTATA | AATCCTCTGA | 1980 |
| 20 | AGGTAGTGCG | ACACTTGATT | TAACGAAAGT | ATATTATGGT | ATGCAAATGC | AAATGATGAC | 2040 |
| • | ATACATGGAT | ATCGTTTTAC | AAAATAAACA | ACGCCTTGGA | TTAACAGATA | TTGTGAAcCA | 2100 |
| • | GGTGGaTTAT | TATACTTCCA | TGTACATGAA | CCTAGAATTA | AATTTAAATC | ATGGTCTGAT | 2160 |
| 25 | ATTGATGAAG | ATAAACTAGA | ACAÄGATTTA | ATTAAAAAGT | TTAAGTTGAG | TGGTTTAGTT | 2220 |
| | AATGCAGACC | AAACTGTTAT | TGATGCATTG | GATATTCGTT | TAGAACCTAA | ATTCACTTCA | 2280 |
| | GATATTGTAC | CAGTTGGTTT | GAATAAAGAT | GGCTCTTTGA | GTAAACGAGG | CAGCCAAGTG | 2340 |
| 3 0 | GCAGATGAAG | CAACGATTTA | TAAATTCATC | CAACATAACA | AAGAGAATTT | TATAGAAACA | 2400 |
| | GCTTCAAATA | TTATGGATGG | ACATACTGAA | GTTGCACCAT | TAAAGTACAA | ACAAAAATTG | 2460 |
| | CCATGTGCTT | TTTGTAGTTA | TCAATCGGTA | TGTCATGTAG | ATGGCATGAT | TGATAGTAAG | 2520 |
| 35 | CGATATCGAA | CTGTAGATGA | AACAATAAAT | CCAATTGAAG | CAATTCAAAA | TATTAACATT | 2580 |
| | aatgātgaat | TTGGGGGTGA | GCAATAGATG | ACAATTCCAG | AGAAACCACA | AGGCGTGATT | 2640 |
| 40 | TGGACTGACG | CGCAATGGCA | AAGTATTTAC | GCAACTGGAC | AAGATGTACT | TGTTGCAGCC | 2700 |
| | GCGGCAGGTT | CAGGTAAAAC | AGCTGTACTA | GTTGAGCGTA | TTATCCAAAA | GATTTTACGT | 2760 |
| | GATGGCATTG | ATGTCGATCG | ACTITTAGTC | GTAACGTTTA | CAAACTTAAG | CGCACGTGAA | 2820 |
| 45 | ATGAAGCATC | GTGTAGACCA | ACGTATTCAA | GAGGCATCGA | TTGCTGATCC | TGCAAATGCA | 2880 |
| | CACTTGAAAA | ACCAACGCAT | CAAAATTCAT | CAAGCACAAA | TATCTACACT | CCATAGTTTT | 2940 |
| | TGCTTGAAAT | TAATTCAACA | GCATTATGAT | GTATTAAATA | TTGACCCGAA | CTTTAGAACA | 3000 |
| 50 | AGCAGTGAAG | CTGAAAATAT | TTTATTATTA | GAACAAACGA | TAGATGAGGT | CATAGAACAA | 3060 |
| . • | CATTACGATA | TCCTTGATCC | TGCTTTTATT | GAATTAACAG | AGCAATTGTC | TTCAGATAGA | 3120 |

| | AATCCTACAA | ATIGGTIGGA | TCAATIGGTG | ACACCATACG | AAGAAGAAGC | ACAACAAGCG | 3240 |
|----|------------|------------|------------|------------|------------|------------|------|
| | CAACTTATTC | AACTACTAAC | AGACTTATCT | AAAGTATTTA | TCACAGCTGC | TTATGATGCT | 3300 |
| 5 | TTAAATAAGG | CGTATGATTT | GTTTAGTATG | ATGGATAGCG | TCGATAAACA | TTTAGCTGTT | 3360 |
| | ATAGAAGATG | AACGACGTTT | AATGGGGCGT | GTTTTAGAAG | GTGGCTTTAT | TGATATACCT | 3420 |
| | TATTTAACTG | GTCACGAATT | TGGCGCGCGT | TTGCCTAATG | TAACAGCGAA | AATTAAAGAA | 3480 |
| 10 | GCAAATGAAA | TGATGGTCGA | TGCCTTAGAA | GATGCTAAAC | TTCAGTATAA | AAATATAAA | 3540 |
| | TCATTAATTG | ATAAAGTGAA | GAGTGATTAC | TTTTCAAGAG | AAGCTGATGA | TTTGAAAGCT | 3600 |
| 15 | GATATGCAAC | AATTGGCGCC | ACGAGTAAAG | TACCTTGCGC | GTATTGTGAA | AGATGTTATG | 3660 |
| .0 | TCAGAATTCA | ATCGAAAAAA | GCGTAGCAAA | AATATTTTGG | ATTTTTCTGA | TTATGAACAT | 3720 |
| | TTTGCATTAC | AAATTTTAAC | TAATGAGGAT | GGTTCGCCTT | CAGAAATTGC | CGAATCATAC | 3780 |
| 20 | CGTCAACACT | TCCAAGAAAT | ATTGGTCGAT | GAGTATCAAG | ATACGAACCG | AGTTCAAGAG | 3840 |
| | AAAATACTAT | CTTGCATCAA | AACGGGTGAT | GAACATAATG | GTAATTTATT | TATGGTTGGA | 3900 |
| | GATGTTAAGC | AATCCATTTA | TAAATTTAGA | CAAGCTGATC | CAAGTTTATT | TATTGAAAAG | 3960 |
| 25 | TATCAACGCT | TTACTATAGA | TGGAGATGGC | ACTGGACGTC | GAATTGATTT | GTCGCAAAAC | 4020 |
| | TTCCGTTCTC | GAAAAGAAGT | ACTGTCAACG | ACTAACTATA | TATTCAAACA | TATGATGGAT | 4080 |
| | GAACAAGTCG | GTGAAGTAAA | ATATGATGAA | GCGGCACAGT | TGTATTATGG | TGCACCATAT | 4140 |
| 30 | GATGAATCGG | ACCATCCAGT | AAACTTAAAA | GTCCTTGTTG | AAGCGGATCA | AGAACATAGT | 4200 |
| | | | | | - | AGATATCTTA | 4260 |
| 35 | | | | | | ATACAAGGAT | 4320 |
| 33 | | | | | | CTTTAAAAAT | 4380 |
| | | | | | | AGAAGTCCGC | 4440 |
| 40 | | | | | | TTTAGTTGGG | 4500 |
| | | | | | TAGCTCAAAT | TAGAATATTG | 4560 |
| | AGTCCAAATG | ATGACTACTT | CTATCAATCG | ATTG | | | 4594 |

(2) INFORMATION FOR SEQ ID NO: 201:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6313 base pairs
- (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear

55

50

| | GGTTTTCnTG | GAAAGATAGT | GAAAATCTCG | TGTTTTTTGG | TTTTGAGGTG | TTGTTTGTAT | 60 |
|----|------------|------------|------------|------------|------------|------------|------|
| | TTTATAAAAT | GGCTTACATA | TATGAAGCGT | TGATTAAGTA | TGGAATTGTT | AATTAATTGA | 120 |
| 5 | ACCTATTTAG | CTTTAAGAAG | GCATAACAAG | ATGACCTTAT | TTTATGCTAT | AATATTTCTA | 180 |
| | TTATGCGAAG | ATTAAGGTGA | GTAGTAAATT | GGATAAAAA | GTAAGTATTC | AAACAAAGCA | 240 |
| | AGTGTTGAAA | CAGCACAACG | AAAAAGAAAA | ATTTGAATTT | ACTACTGAAG | GAACTTGGCA | 300 |
| 10 | ACAAAGGCAA | TCTAACTTTA | TTCGGTATGT | AGAACAAATT | GAGGATGCAA | CAGTTAATGT | 360 |
| | TACAATAAAA | GTGGATGATG | ATAGCGTTAA | GTTGATTCGT | AAAGGCGACA | TTAATATGAA | 420 |
| 15 | TTTGCATTTT | GTTGAAGGAC | AAACGACAAC | AACTTTTTAC | GATATATCGG | CTGGACGAAT | 480 |
| | TCCACTAGAA | GTTAAAACAT | TACGCATTTT | ACATTTCGTA | AGTGGAGACG | GTGGCAAGCT | 540 |
| | AAAGATTCAT | TATGAATTAT | ATCAAGATAA | TGAAAAAATG | GGTTCTTATC | AATATGAAAT | 600 |
| 20 | TAACTATAAG | GAGATAGGCG | AATGAATATT | ATTGATCAAG | TGAAACAAAC | ATTAGTAGAA | 660 |
| | GAAATTGCAG | CAAGTATTAA | CAAAGCAGGA | TTAGCAGATG | AGATTCCTGA | TATTAAAATT | 720 |
| | GAAGTTCCTA | AAGATACAAA | AAATGGAGAT | TATGCTACTA | ATATTGCGAT | GGTACTGACT | 780 |
| 25 | AAGATTGCAA | AGCGTAATCC | TCGTGAAATT | GCTCAAGCGA | TTGTTGATAA | CTTAGATACT | 840 |
| | GAAAAAGCAC | ATGTAAAACA | AATTGACATT | GCTGGTCCAG | GATTCATTAA | TTTTTACTTA | 900 |
| | GATAATCAGT | ATTTAACAGC | AATTATTCCT | GAAGCAATTG | AAAAAGGTGA | TCAATTTGGA | 960 |
| 30 | CATGTAAATG | AATCAAAAGG | TCAAAATGTA | TTGCTTGAGT | ATGTTTCAGC | TAACCCTACA | 1020 |
| | GGAGATTTAC | ATATTGGTCA | TGCTAGAAAT | GCAGCAGTTG | GTGATGCTTT | AgcTAAtATT | 1080 |
| 35 | TTAACTGCAG | CTGGCTATAA | TGTAACACGT | GAATATTATA | TTAATGATGC | TGGTAATCAA | 1140 |
| ,, | ATTACTAACT | TAGCGCGTTC | GATTGAAACA | CGTTTCTTTG | AAGCTTTAGG | TGACAATAGT | 1200 |
| | TATTCAATGC | CAGAAGATGG | CTATAATGGA | AAAGATATTA | TTGAAATAGG | TAAAGATTTA | 1260 |
| 10 | GCAGAGAAAC | ACCCTGAAAT | TAAAGATTAT | TCTGAAGAAG | CACGTTTGAA | AGAATTTAGA | 1320 |
| | AAATTAGGCG | TAGAATACGA | AATGGCTAAA | TTGAAAAATG | ATTTAGCAGA | GTTCAATACG | 1380 |
| | CATTTTGATA | ATTGGTTTAG | TGAAaCATCT | TTATATGAAA | AAGGAGAAAT | TCTTGAAGTT | 1440 |
| 15 | TTAGCAAAAA | TGAAAGAATT | AGGTTATACG | TATGAAGCTG | ATGGCGCTAC | ATGGTTACGT | 1500 |
| | ACAACTGATT | TTAAAGACGA | CAAAGACAGA | GTATTAATTA | AAAATGACGG | TACATATACG | 1560 |
| | TATTTCTTAC | CAGATATTGC | GTACCÁCTTC | GATAAAGTAA | AACGTGGTAA | TGACATTTTA | 1620 |
| 50 | ATCGATTTAT | TTGGTGCTGA | TCATCATGGT | TATATTAATC | GTTTGAAAGC | ATCTCTTGAA | 1680 |
| | ACGTTTGGTG | TAGATAGTAA | TCGTTTAGAA | ATTCAAATCA | TGCAAATGGT | TCGTTTAATC | 1740 |

| | ATTATGGACG | AAGTTGGCGT | TGACGCTGCA | CGTTATTTCT | TAACTATGCG | TAGTCCTGAT | 1860 |
|----|------------|------------|------------|------------|------------|------------|------|
| | AGTCACTTTG | ATTTTGATAT | GGAATTAGCG | AAAGAGCAAT | CTCAAGACAA | TCCAGTTTAC | 1920 |
| 5 | TATGCTCAAT | ATGCACATGC | GCGTATTTGT | TCAATTTTAA | AACAAGCGAA | AGAGCAAGGT | 1980 |
| | ATTGAAGTGA | CTGCTGCGAA | TGATTTTACA | ACGATTACTA | ATGAAAAAGC | GATTGAATTG | 2040 |
| | TTGAAAAAAG | TAGCTGATTT | CGAACCTACA | ATTGAAAGTG | CTGCTGAGCA | TAGATCGGCA | 2100 |
| 0 | CATAGAATTA | CTAATTATAT | TCAAGATTTA | GCTTCTCATT | TCCATAAATT | CTATAATGCT | 2160 |
| | GAAAAAGTGT | TAACAGATGA | TATTGAAAAA | ACAAAAGCAC | ATGTTGCTAT | GATTGAAGCG | 2220 |
| | GTCAGAATTA | CATTGAAAAA | TGCATTGGCA | ATGGTCGGTG | TAAGCGCACC | TGAATCAATG | 2280 |
| 5 | TAAGAACATT | TATATACACT | CCAACGTAGA | GTTTCTCGAA | AGATACTTTG | TGTTGGAGTG | 2340 |
| | TTTTTTTAG | GTATGTGACA | TATTGGGGAA | TGCTTAGTAT | GTGAATAAGG | TTAAGAGGAA | 2400 |
| 20 | CACAGTTGGA | TGCTCTGCAC | AACTGCATAA | GAGAGCCTGA | GACATAAATC | AATGTTCTAT | 2460 |
| .0 | GCTCTACAAA | GTTATAATGG | CAGTAGTTGA | CTGAACGAAA | ATTCGCTTGT | AACAAGCTTT | 2520 |
| | TTTCAATTCT | AGTCAACCTT | GCCGGCGGG | CCCCAACAAA | GAGAAATTGG | ATTCCCAATT | 2580 |
| | TCTACAGACA | ATGCAAGTTG | GGGTGGGACG | ACGAAATAAA | TTTTACGATA | ATATCATTTC | 2640 |
| | TGTCCCACTC | CCTCTAAAAT | GGAGGGTGTA | AATGTTAGGA | ACTGATGAAT | TATATAAAGT | 2700 |
| | TTTATATGAA | CATCTCGGAC | CACAATTTTG | GTGGCCTGCT | GATAATGACA | TTGAAATGAT | 2760 |
| 30 | GTTAGGTGCA | ATTTTAGTTC | AAAATACTAG | ATGGCGAAAT | GCAGAAATTG | CATTGAATCA | 2820 |
| | GATTAAAGAA | CATACGCATT | TTAATCCAAA | TCATATATTA | GAACTACCTA | TTGAAACGTT | 2880 |
| | ACAATCATTG | ATACATTCAA | GTGGCTTTTA | TAAAAGTAAA | TCACTGACGA | TTAAAACATT | 2940 |
| 35 | ATTAACATGG | TTAGCACGAC | ATCATTTCAA | TTATCAAGAG | ATTAATGAGC | GATATAAAGG | 3000 |
| | TGGÃTTAAGA | AAAGAATTAT | TATCTTTGAA | AGGTATTGGA | AGTGAAACAG | CAGATGTCTT | 3060 |
| | ACTIGITIAT | ATATTCGGAC | GTATTGAATT | TATTCCAGAT | AGCTATACAA | GAAAAATATA | 3120 |
| 40 | TGATAAATTA | GGATATGAAA | ACACTAAAAA | TTATGATCAA | TTAAAAAAAG | TAGTCACATT | 318 |
| | ACCAAATCAT | TTTACAAATC | AAGATGCTAA | TGAATTTCAT | GCTCTGTTAG | ATGTATTTGG | 324 |
| 46 | TAAACATTAC | TTTAGAGACA | AAGATATAAA | GAATTATGAT | TTTTTAGAAC | CTTACTTTAA | 330 |
| 45 | AAAGTAAACG | CTGTGAAGTT | AGATAGATGA | GTTTATATGA | AAAAATATAA | ATAATTTACT | 336 |
| | | | | | | AAAAACAGTT | 342 |
| 50 | | | | | | GGGAGTAAAG | 348 |
| • | a | | | | | CTCAACACA | 354 |

| | AACGCAGTTG | GATGCTACCG | CACAACTGCA | TAAATCCCTC | TaATCgcTAA | AGCGAAAAGT | 366 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | GGGATTAAAA | AGGAGATGTG | ATAGTGTGAA | GAAATCGTTA | ATTGCTTTTA | TTTTGATTTT | 372 |
| 5 | TATGCTTGTC | CTGAGTGGCT | GTGGTATGAA | AGATAATGAT | AAACAAGGTA | GCAATGATAA | 378 |
| | TGGCTCGTCT | AAATCGCCGT | ACCATAGAAT | TGTTTCGTTA | ATGCCTAGTA | ATACTGAAAT | 384 |
| | TTTATATGAA | TTAGGATTAG | GTAAATACAT | AGTTGGTGTT | TCAACGGTTG | ATGATTATCC | 390 |
| 10 | AAAAGATGTG | AAAAAGGGTA | AGAAACAATT | TGATGCTTTG | AATCTAAATA | AAGAGGAACT | 396 |
| | | | | GCATGAGTCG | | | 402 |
| | | | | CAAAGTAGTG | | | 4080 |
| 15 | | | | AATTGGGAAA | | | 4140 |
| | | | | TAATATCGAT | | | 4200 |
| 20 | | | | TGAGGTTTCA | | | 4260 |
| 20 | | | | GTTAGAAAAA | | | 4320 |
| | | | | GAAGGAAAGT | | | 4380 |
| 25 | | | | ATCAGATTAT | | | 4440 |
| | | | | GAATACACGT | • | | 4500 |
| | | | | | | GAGATGCAAT | 4560 |
| 3 0 | | | | CCTTATTGCT | | | 4620 |
| | | | | | | ATGTGATTGC | 4680 |
| | | | | GAGAATGCTA | | | 4740 |
| 35 | | | • | ACTAGCATAT | | | 4800 |
| | | | | TTAATCAATG | | | 4860 |
| | • | | | GTTGCAGGCC | | | 4920 |
| 10 | | | | GGATTGGCAA | | | 4980 |
| | ودي | | | TTATGGATTC | | | 5040 |
| | | | | | | AGGCTATCCA | 5100 |
| 15 | | | | | | ACTTCTATAT | 5160 |
| | | | | AATACAATTG | | | 5220 |
| 50 | | | | | | CATTGCATTG | |
| _ | | | | | | ACTAAAAAGT | 5280 |
| | | | | | | | |

| | ATAACGCCGA | TAAATGTCGC | ATATGTTGGC | ATCATTGGAT | TCATTGGTAT | GGTGATACCG | 5460 |
|----|------------|------------|------------|------------|------------|------------|------|
| | CAACTCATTA | GAAAATGGCA | GTGGAAACAA | TCATTAGGAA | GACAATTGGC | TTTAAATATT | 5520 |
| 5 | GTAACTGGAG | GACAAATAAT | GGTTATGGCA | GATTTTATTG | GTAGCCATAT | ATTGTCACCA | 5580 |
| | GTACAAATAC | CGGCAAGTAT | TATCATTGCA | TTAATTGGTA | TACCAGTGTT | AtTTTACaTG | 5640 |
| | CkAAwAtCtC | aGTCgAAAcG | GTTACaCTAG | CACACGACaT | TTGCTAAAAT | AAAAATAACT | 5700 |
| 10 | ATAAACATAA | AGAGGGCATA | AGCGATGGAT | TTGAATCAAA | TTAAAGCAGT | TGTATTTGAT | 5760 |
| | TTAGAAGGTA | CGTTGTTGGA | CAGAGTTAAA | TCTCGAGAGA | AATTTATCGA | AGAGCAATAT | 5820 |
| 15 | GAACGATTTC | ATGACTACTT | AATTCATGTT | CAACTGGCAG | AAAAATTTTA | AgCATTTATT | 5880 |
| 15 | GAGCTAGATG | ACGATGAAGA | TAATGATAAA | CCTGATTTAT | ATAAAGAAAT | CATTAAACGT | 5940 |
| | TTCCATGTAG | ATAGGTTAAC | TTGGAAAGAC | TTATTTAATG | ATTTTGAAAT | GCATTTTTAT | 6000 |
| 20 | CGTTATGTAT | TTCCTTATTA | CGATACTTTG | TATACACTAG | AAAAgCTATC | GCAAAAAGGC | 6060 |
| | TTTCAAATTG | GTGTTATCGC | AAATGGTAAA | TCTAAGATTA | AACAATTTCG | ATTACATTCA | 6120 |
| | CTTGGTTTGA | TGCATGTTAT | TAATTATTTA | TCAACATCAG | AAACAGTTGG | TTTTCGTAAA | 6180 |
| 25 | CCACATCCTA | AAATTTTTGA | AGATATGATT | GATCAACTAG | GGGTATTACC | TGAGCAAATT | 6240 |
| | ATGTATGTTG | GCGATGATGC | GTTAAATGAT | GTAGCTCCAG | CACGAGCTAT | GGGCATGGTT | 6300 |
| | AGTGTATGGT | ATA | | | | | 6313 |

(2) INFORMATION FOR SEQ ID NO: 202:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2174 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 202:

CCGTAAACAC ATCAACAAAA GAAGGCTATA TTACAAAAGA AGACTTGGAC TTATGCTGCA 60 CGTCGCTCTA ATTCAGCTGG AATGCAAGTC ACCGGACGAC TGGCTTACAT TGAACCTTAT 120 GGGGCAACAA GTCGCACAAA ATAAACGCGC GAGAAGCAAG AATAGGAAGT GATATCTATG 180 AAATGGTTAT CACGAATATT AACAGTAATA GTGACCATGT CLATGGGGTG TGGTGCATTG 240 ATATTTAATC GTAGACATCA GCTAAAGGCG AAAACGCTGA ACTTCAATCA TAAAGCATTA 300 ACANTTATTA TTCCGGCTAG AAACGAAGAA AAAAGAATAG GTCATTTACT ACATTCGATA 360 ATACAACAGC AAGTTCCAGT AGATGTCATT GTTATGAATG ACGGATCGAC AGATGAAACA 420

55

35

40

45

50

| | AAATGGTATG | GGAAATCACA | TGCTTGTTAT | CAAGGTGTGA | CGCATGCATG | TACGAATCGC | 540 |
|----|------------|------------|------------|------------|------------|------------|------|
| | ATTGCCTTTG | TAGATGCTGA | TGTAACTTTC | TTAAGGAAAG | ATGCTGTTGA | AACGTTGATT | 600 |
| 5 | AATCAGTATC | AATTACAAGG | TGAAAAAGGA | TTGTTAAGCG | TACAGCCTTA | TCATATAACA | 660 |
| | AAGCGTTTCT | ACGAAGGGTT | TTCAGCGATA | TTTAATTTAA | TGACAGTCGT | TGGTATGAAT | 720 |
| •• | GTATTTTCTA | CCTTAGACGA | CGGTCGGACT | AACCAGCATG | CATTTGGACC | GGTGACATTA | 780 |
| 10 | ACAAATAAAG | AAGATTATTA | TGCAACTGGA | GGTCATAAAA | GTGCAAACCG | TCATATTATT | 840 |
| | GAAGGATTTG | CTTTAGGAAG | TGCATATACT | TCACAATCAT | TGCCCGTAAC | AGTTTATGAA | 900 |
| 15 | GGGTTTCCAT | TTGTTGCATT | TCGCATGTAT | CAAGAAGGAT | TTCAGTCATT | ACAAGAAGGA | 960 |
| | TGGACAAAGC | ATTTGTCAAC | TGGGGCAGGT | GGCACAAAGC | CTAAGATCAT | GACAGCAATT | 1020 |
| | GTGTTGTGGT | TGTTTGGTTC | TATAGCGAGT | ATTTTAGGGC | TATGTCTTAG | TTTAAAATAT | 1080 |
| 20 | CGCCAAATGT | CTGTAAGAAA | AATGGTAGCA | CTTTACTTGA | GCTATACTAC | ACAATTTATT | 1140 |
| | TATCTGCATC | GAAGGGTCGG | CCAATTTTCT | AATTTATTAA | TGGTATGTCA | TCCATTGTTA | 1200 |
| | TTTATGTTTT | TTACTAAAAT | TTTCATCCAA | TCTTGGAAAC | AAACGCATCG | TTATGGTGTA | 1260 |
| 25 | GTTGAATGGA | AAGGTCGTCA | ATATTCTATA | TCTAAAGAAC | AATAAATCAA | GGTAATGGCA | 1320 |
| | TTTCAATATA | GGAGGACTAG | TATGACAATG | ATGGATATGA | ATTTTAAATA | TTGTCATAAA | 1380 |
| | ATCATGAAGA | AACATTCAAA | AAGCTTTTCT | TACGCTTTTG | ACTTGTTACC | AGAAGATCAA | 1440 |
| 30 | AGAAAAGCGG | TTTGGGCAAT | TTATGCTGTG | TGTCGTAAAA | TTGATGACAG | TATAGATGTT | 1500 |
| | TATGGCGATA | TTCAATTTTT | AAATCAAATA | AAAGAAGATA | TACAATCTAT | TGAAAAATAC | 1560 |
| | CCATATGAAC | ATCATCACTT | TCAAAGTGAT | CGTAGAATCA | TGATGGCGCT | TCAgCATGTT | 1620 |
| 35 | GCACAACATA | AAAATATCGC | CTTTCAATCT | TTTTATAATC | TCATTGATAC | TGTATATAAA | 1680 |
| | GATCAACATT | TTACAATGTT | TGAAACGGAC | GCTGAATTAT | TCGGATATTG | TTATGGTGTT | 1740 |
| 10 | GCTGGTACAg | TAGGTGAAGT | ATTGACGCCG | ATTTTAAGTG | ATCATGAAAC | ACATCAGACA | 1800 |
| , | TACGATGTCG | CAAGAAGACT | TGGTGAATCG | TTGCAATTGA | TTAATATATT | AAGAGATGTC | 1860 |
| | GGTGAAGATT | TTGACAATGA | ACGGATATAT | TTTAGTAAGC | AACGATTAAA | GCAATATGAA | 1920 |
| 15 | GTTGATATTG | CTGAAGTGTA | CCAAAATGGT | GTTAATAATC | ATTATATTGA | CTTATGGGAA | 1980 |
| | TATTATGCAG | CTATCGCAGA | AAAAGATTTT | CAAGATGTTA | TGGATCAAAT | CAAAGTATTT | 2040 |
| | AGTATTGAAG | CACAACCAAT | CATAGAATTA | GCAGCACGTA | TATATATTGA | AATACTGGAC | 2100 |
| 50 | GAaGTGAGaC | AGGCTAACTA | TACATTACAT | GAACGTGTTT | TTGTGGaTAA | GAGGAAAAAG | 2160 |
| | GCAAAGTTGT | TTCA | | | | | 2174 |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4715 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 203:

| 4.0 | | | | | | | |
|-----|------------|------------|------------|------------|------------|------------|------|
| 10 | GAAnCAGnTA | GACAAATTAT | GGaAAmCGGT | GTGAATCaAG | GATTCtTTGG | TGTAGCTGGT | 60 |
| | TTTGACCTAC | TCGTCGATGA | GGATGATAAC | GTTTATGCGA | TTGATTTAAA | CTTTAGACAA | 120 |
| 15 | AATGGTTCaA | CGAGCATGTT | ATTACTTGCT | AACGAGTTGA | ATTCAGGATA | TCAAAAGTTT | 180 |
| | TATAGTTATC | ATTCAAAAGG | TGATAACACA | CATTTCTTCA | ATACGATTTT | GAAATATGTC | 240 |
| | AAAGAAGGTA | GTTTATACCC | GTTATCTTAT | TATGATGGTG | ATTGGTACGG | TGAAGATAAA | 300 |
| 20 | GTTAAATCAA | GGTTTGGCTG | TATTTGGCAT | GGTGATTCAA | AAGAAACAGT | ACTGGAGAAT | 360 |
| | GAACGCGCAT | TTTTAGCTGA | ACTTGAACAC | TATTAGAGTT | CGGAACATAA | GGCGCTACAA | 420 |
| | TGTTGTGTTG | CCAGTAGTTG | ACTGAATATG | CGTTTGTAAC | AAGCTTTTTT | CGATTCTAGT | 480° |
| 25 | CAACAGTAAT | TAAATTTATG | ATATGGCAAT | ACTTTGTAAT | ACTAATATTA | AATGGCGACT | 540 |
| | TTTATTTCAC | TATGTTATAA | GAGTTGCCAT | TTTGTTGATA | AAGGTATACT | AAAGGTTATC | 600 |
| | GTTTTGAAAT | TTTTAGTAAC | TAGATATGTT | TCGTGTTATA | GACCGAATTT | GTGTATACGT | 660 |
| 30 | AAAATTTAAT | GCTATTGAAT | TTTTAAAATG | AAAAACATGA | CATTAAATTG | AATTCATAAT | 720 |
| | ATGTCTAATT | GACTAACTTG | TTGGAGTCAT | TTACTATTTT | ATGTATGACA | TATTTTAAAA | 780 |
| | AGTGAGGGTC | AAGCATGTCT | TATAAAGCAT | ATCCATTCTT | TAGAGATATA | TTAATAAATG | 840 |
| 35 | AATGTATTTA | TTTCGCCTCT | AAAATAAAA | AACTAGTACG | CCTAAATTAT | AAAAGTGAAG | 900 |
| | CGnATGTAGG | CGTTTGGACA | GAAGAAAGTG | TGGCCGTATC | ATTTTTAACA | AGTCGTGATA | 960 |
| 10 | TTCCATTTGA | TAAAGTTGTA | AAAATGGACG | TTGATCGTTT | TGCTACTTAT | GAATTAGATG | 1020 |
| •0 | AATTGTTTGA | TGAACAAGAC | CATATTATTA | TGAATCAAAC | AATGGAAGAW | GAAGGGCATC | 1080 |
| | TACTAAACGT | TGTAGCTGTT | ACACAAGAAG | TGATGACGGA | ATTAGATAAA | ATTAGAATCA | 1140 |
| 15 | AAGAATTTGT | CCAAGATGTA | GCGAAATATG | ATGAAGTATA | CGGCTTAACT | AAAAAAGGTA | 1200 |
| | GTAAGCAGTT | TATTCTCATT | AGTGAAAATG | ATAGCGACGA | AAAAAAGCCG | CATATTATGC | 1260 |
| | CTGTATGGAG | TATTAAAAAC | AGAGCGTTAA | AAGTTCGAGA | TGAAGATTTT | GAAGAGTGTG | 1320 |
| 50 | ATTTAATTAC | GATTGAAGGT | TCTGTTTTCG | GAGAATGGCT | AGATGAACTT | AGAGATGATC | 1380 |
| | ATAAAGCCGT | TGCGATAGAT | TTAAAAACTG | GCGTGGTTGG | TACAATTGTT | TCAGCGCAAA | 1440 |

55

| | ATGGAACAA: | I ACGTATTCAA | AACACTTAGA | CCATAAAATA | AAAGGCCATT | TATATAGCGT | 1560 |
|------------|------------|--------------|------------|------------|------------|------------|-------|
| | TTATTTAAA | A CAACGCGCAT | ATAAATGGTC | TTTTTCTATT | TTTCTAAATA | TAATGCACCA | 1620 |
| 5 | ATAGCACCTO | naaaatgcgc | CGTTTTCAAC | ATAGTACGGT | TTGCAACCGC | GTAACACAGT | 1680 |
| | ATAATCTTC | C ACAACTTTGC | GTAATAAAGC | GTTATTATGA | AATGAAGAAC | CGATATAAAC | 1740 |
| 10 | GATATTTTC | GTTTTAAATT | CACGTGCAAC | AGTAATGGCC | ATTGTCGTAA | CAACTTCGCC | 1800 |
| 10 | AACGACACCA | ATAACGGCTG | CTAATTTATT | GCTAGGTGTA | AAATCAGCAT | CTAAATGATG | 1860 |
| | TAGTACATGA | CCAAAATTAG | CTGCTGTTAA | ATCACCGGGA | ATGGGTGGTT | CGGTATCTTT | 1920 |
| 15 | ATAAATATGT | CTAACCTTTA | AATCGATAGT | GTTACGATCA | CCGTGTTGTG | CCATGTCAGT | 1980 |
| | TAACTGTTTA | TAATCAGTGA | TTTGACTTAG | TAAATAACCG | AGTCCTTGAA | TCATGCCTCC | 2040 |
| | ACCTGTACCG | ATACCGCCTA | CACGACGTTG | TGATTGGCCG | TCGAAATAAT | GTAGTGACGT | 2100 |
| 20 | ACCGGTACCA | ACATTTGCAA | AAATATAATC | TGCTAAGTCA | TGGCCTTGCT | CTTTTAACAA | 2160 |
| | AATACCTAGT | CCTTGAGATG | CAGCATCAAA | CTCTACAAAA | ATTTGTGCAG | GAATGTTGAT | 2220 |
| | GTTTTCAGCA | ATGACACCTG | CATTACCTCC | AGTTAAGCAT | AATTTTTCAA | TTTGCTGTTG | 2280- |
| 25 | GTTTAACCAT | TCCACAACTT | GATCAATATT | TTTAGTTAAT | TCAGTTTTAA | AAGTACGTTG | 2340 |
| | GTTATCTTGC | TCTTGAACGA | TTTTAATTAG | TGTACCGCCA | GCGTCAATGC | CAACTTTCAT | 2400 |
| | AAGATTCCCA | CCTCATTATT | AATGTCTATC | CTTAAATAAT | AGTATAGTAA | AATGACTAAA | 2460 |
| 30 | AAACAAGTAA | TAATAGTAAT | TATTAACAAA | TTTGATGCCa | TTGCATTTCA | ACATTGTAAG | 2520 |
| | CGTATCGCAA | TTAAtGTTTT | ACAAACGTGG | ACGTTAAGTt | ATATATATTA | TTTTCTAGGA | 2580 |
| | ATTTTGAAGT | TGTATAGGAT | TGTTAGTTAG | TGACGCAATA | TTAAAAGTAG | TTCGTACGCA | 2640 |
| 35 | GTGTATTTGT | AAGTCTCTGA | TTAAAATGAT | AAGTAATGAG | GAATAGTACA | TTAATTTTGA | 2700 |
| | AATŢTAAAAA | ATATAAATAA | GTAATTTATT | TAACTTAGAG | CAAATAATGG | TATCGTAGTG | 2760 |
| 10 | AAATAATAGG | TAAAATAATA | TGGGGATTCA | TGCTTCATAT | ATAAAAAGAT | AGGGGTTAAA | 2820 |
| | TATATGGCTA | AAGAACTTTG | TTTTGAAGGT | ATCACTTTAA | AAGCATTTGA | TGAACAATAT | 2880 |
| | CGTTCAGCAA | TTAATGATTT | TGACTTGAAT | GAAAGACAAC | AAATATATTC | ATCTTTACCT | 2940 |
| 1 5 | AAAGAAGTTA | TTGATGATGC | AATTAATGAT | GCTGATAGGA | TTGCTAACGT | AGCAWTAAmC | 3000 |
| | GATAAAAATG | AAGTGGTGGG | CTTTTTTGTA | TTACATCGTT | ACTATCAGCA | TGAAGGTTAT | 3060 |
| | GATACACCTG | AAAATGTCGT | TTATATTCGT | TCATTATCGA | TTAATGAAAA | ATATCAAGGT | 3120 |
| 50 | TTTGGATATG | GCACGAAAAT | AATGATGTCA | TTGCCGCAAT | ATGTTCAAGG | TGTATTTCCT | 3180 |
| | GATTTTAATC | ATCTATATCT | AGTAGTAGAT | GCGGAAAATG | ACAATGCTTG | GAACCTATAC | 3240 |

| | CTATATTACT | TGGACTTAGA | TTCAAAACAT | GTTTCATCAT | TAAAGCTTGA | AGAAGAAAGT | 3360 |
|-----|------------|------------|------------|------------|------------|------------|------|
| | CGTTCAGAAG | TGACCAATGT | ACATATCATT | AATTTAATGA | TTGATGGCCA | AAAGGTTGGC | 3420 |
| 5 | TTTATCGCAT | TGGAGCAGAT | TGGTGAACGC | ATGAACATTG | CTGCTATTGA | AGTGGATAAA | 3480 |
| | TCATATCGCT | TTAATGGTAT | TGGTTCAAGT | GCTCTGCGAC | AATTGCCAAC | TTACTTAAGA | 3540 |
| 10 | AAAAACTATG | ACAACCTTAA | TGTGATTACG | ATGATTCTGT | TTGGAGAGAA | TAATGATTTT | 3600 |
| , 0 | AAACCATTAT | GTTTAAATAG | TAATTTCGTT | GAAATCGAAC | AAACTGATGA | TTATGTCGTT | 3660 |
| | TTCGAAAAAT | ATTTAAATTA | CTAACAGTGA | TTGCGAAATA | TGATATTGTC | ATTTATAATT | 3720 |
| 15 | TAGTTTTGTT | ACTATATATA | AATGAATTCA | GACGTATAAA | TTTAGATTAT | ATCCTTCGAA | 3780 |
| | AGGAAGTATT | GGGCAATGAA | AATTCAAGAT | TATACAAAAC | AAATGGTTGA | TGAAAAATCA | 3840 |
| | TTTATTGATA | TGGCTTATAC | ATTATTGAAT | GATAAAGGCG | AAACAATGAm | mTTATATGAT | 3900 |
| 20 | ATYATCGATG | AATTTAGAGC | GTTAGGTGAT | TATGAGTACG | AAGAAATTGA | AAATCGTGTT | 3960 |
| | GTACAATTTT | ACACGGATTT | AAACACAGAT | GGTCGTTTTT | TAAATGTTGG | AGAAATTTA | 4020 |
| | TGGGGATTAC | GTGATTGGTA | TTCGGTAGAT | GATATTGAAG | AGAAAATCGC | ACCAACTATT | 4080 |
| 25 | CAAAAATTCG | ATATTCTGGA | TGCAGATGAT | GAAGAAGATC | AAAACTTAAA | ATTATTGGGC | 4140 |
| | GAAGATGAAA | TGGATGACGA | CGATGATATT | CCAGCTCAAA | CAGATGATCA | AGAAGAACTA | 4200 |
| | AATGATCCAG | AAGATGAGCA | GGTTGAAGAA | GAAATCAATC | ATTCGGATAT | AGTCATTGAA | 4260 |
| 30 | GAAGATGAAG | ATGAACTAGA | CGAAGACGAA | GAAGTGTTTG | AAGACGAAGA | AGACTTCAAC | 4320 |
| | GATTAATTTT | TTGTTTGACT | TTTAGTTGAA | AGATGATAAA | ATTTTATTCG | GGCTCCTTTA | 4380 |
| 35 | AATAGGACAC | GTGTATAAAA | TTTATACGCT | CCCCTTACAG | AATTTGTGAG | AGGGAGCGTT | 4440 |
| 00 | TTTTTATTTA | ATTGAGTAAA | TCAAGAAATG | ATAACGCAAA | AATCAAAGTT | GTAAATGATA | 4500 |
| | TACATAGTGA | CATAGCAGTA | TGGAAACGGT | AAGTAAACAG | AATTTAATTT | TGTCGALTCG | 4560 |
| 40 | ACAATAAACA | aCTtGAaTGA | GCTTGCTTTA | ATGTTATGTn | nTACGTAATT | TTTACAATTG | 4620 |
| | ATGAGGAAGC | ATTCCCTTTA | ATAATTAGGA | GGTCAAGACA | TGACAAAATT | TATTTTTGTA | 4680 |
| | ACAGGTGGCG | TAGTTTCATC | CATTAGGGGA | AGGGT | | | 4715 |

(2) INFORMATION FOR SEQ ID NO: 204:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 918 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear

55

45

| | | _ |
|----|---|-----|
| | ATAATAACTG AAATTAAAAT TGCTAAATMG TGTTAAGCTA TCGCMACAAT GAAAATWCCG | 60 |
| _ | ATTTTGCGTT GTTGAAAATA TCTTTCCAAA CCAAGAATCG ATAATGGCAA TAAATATAAT | 120 |
| 5 | AAATTTCCAT AAAATGACCA AGTAAAATTA AAGTATATAA CGACAGTTGA CATGCCGTAT | 180 |
| | AAAATCGTAG CGATCATATT TGCTGAGCGT TTAAAGTGTA ATATTTTAAA TAAGTAGAAG | 240 |
| | GTCACGACAA ATGTTATGAT AGCTCGTATC ATAGCCATAA TAAGTTGGTT TGTCGGCCAA | 300 |
| 10 | AAATGTATTG TCGTCGGATT AAATATACCA ACCGTTTCTC CTATTTTAAT GAAKAGAAAA | 360 |
| | TTTAGCCACA TTAAAGGTGA CAGCGAATAA TAATnTGATA GTCCTTTCAT ATAATCGCCA | 420 |
| | CCTAMTCCAA ACGATGCATC ATTTAAACTA GAANAACTAC GTAGATGTTC ATACANATAC | |
| 15 | | 480 |
| | ATTTGAAATG GCATCATTTG ACGGAATCCA TCTCCAGCCC CGCTAAAAAC AGTACCATTC | 540 |
| | ACAATATAAT CATAGATATG AGTAGAAAAT AAAATAAGCG TTAATATTAC ACTAATGAAA | 600 |
| 20 | GTTATAACAA AGAATTGTTT GACGTTTGAA TTTAGCCACT TTTTTAACAC AACATTATCC | 660 |
| | TCAACTTTCA AATTTAAAAT TAAGTTTAAC TGAAACTAAA GTTAATGAGG TTCTTGATAG | 720 |
| | GTAAAGACGA AGATGACTGT GGAACAGATA CCTTATCATA GTTACTTAAA CTTTGGATCA | 780 |
| 25 | TITTCAGTTT ATCATTAAAC AAATATATTG AATAATAAAA ATGTCATACT GATAAAGATG | 840 |
| | AATGTCACTT AATAAGTAAC TTAGATTTAA CAAATGATGA TTTTTAATTG TAGAAAACTT | 900 |
| | GAAATAATCA CKTATACC | 918 |
| 30 | (2) INFORMATION FOR SEQ ID NO: 205: | 710 |
| | | |
| | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 16397 base pairs | |
| 35 | (B) TYPE: nucleic acid (C) STRANDEDNESS: double | |
| 55 | (D) TOPOLOGY: linear | |
| | | |
| | . (vi) CEOUENCE DECODITION OF THE CO. | |
| 40 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 205: | |
| | TCGCCCHATA ATCAATTTAT TTTTCATGTG CCACTCCTAT ACAAGCTHAC AATGCTTCTT | 60 |
| | CAGTTAAGGC AATATCTTTT AATTTTGTTT GATATTTTTG TTCAAAGTCA TATTGTAACT | 120 |

GAACAATTTC TGGCAAACCA ATATGCCAAT CCGCCAATTT TTTTTTAYCT TtGAAGAGCT

CTTTTGGTGA TGKTTGCGAC ACTATACTAC CTTCTTTCAT AACGATGACT TCATCTGCAT

AACGCGCGAC TTCATTCATA TCATGTGAAA TTAGGATAAT TGCCTTATTT TCATCTGTTT

GTAGTGACTT TAGTAATCTC ATTACTTGTC GTTTACTTTG TGGATCAAGT CCTGCTGTAG

GTTCATCAAC CACGATAATA TCAGGATTCA TTGCCAATAT CGATACAATC GCTATTTTAC

180

240

300

360

420

55

50

| | AATCCATCAA | CAGACGATGG | GCATAGTTTT | TGGCTTCATC | TAAATTCATT | TTAAAGTTTT | 540 |
|----|--------------|------------|---|-------------|--------------------------------|----------------|--------|
| | TAGGTCCAAA | TATCATTTCA | CGCTCTACTG | TGTCCTCAAA | TAATTGAGAT | TCGGGAAATT | . 600 |
| 5 | GAAATACCAT | TCCAATTCTT | TTTCTTACAG | GTCTAATATA | TTTATCTTTG | GTCTTATGTG | 660 |
| | TAATAGTAAT | GTCATCAACT | GTAACTGTCC | CAGTAGTCGG | CTTTAACAGC | GCATTAATAT | 720 |
| | TTTGTATCAA | CGTTGATTTA | CCACTACCCG | TTTGTCCAAC | GATGGCGTAA | TATTTACCTT | 780 |
| 10 | GTTCAAATTC | TGTATTAACA | TCATGAATAG | CTTGATGCTG | ATATGGTGTC | CCTTTTTGAT | 840 |
| | AGGTATAACT | TACATTGTCA | AACCGTATAG | TCATAGTTGA | TCCACCAGCC | CTTCATAAGT | 900 |
| 15 | TAAGAATGAT | GTTTGGTGTC | CCAGCATTTG | ATTTATTTTG | ATTGGGAATG | GCAAATCTAG | 960 |
| | ACCTATTCTT | GTTAACTCTT | CTGCATTGTC | GAAAATTTCA | GTCGCTGTGC | CTTCTTTATA | 1020 |
| | GACAGTCCCT | TTATTCATAA | CGATAACATG | ATCTGCTTCC | ATCGCCTCAG | ATAAATCATG | 1080 |
| 20 | CGTAATAGAA | ATGATTGTAA | TATTATGTTC | TGATTTAACT | TTTCTCACTA | AATCCAATAA | 1140 |
| | ATTTTGACGT | GCATCAGGAT | CTAACATAGA | AGTCGCCTCA | TCTAATATAA | TGACAGAGGG | 1200 |
| | GTTAAGTGCT | AATACACTTG | CTATAGCCAC | ACGCTGCTTC | TGTCCCCCG | ATAATGCATT | 1260 |
| 25 | AGGTTCATAA | TCTGCACGTT | CTAACATATC | AACTTGTTTA | AGTGCTTCGC | TGACTCTTCT | 1320 |
| | ATGCATTTCG | TCATATGGAA | CCGCATGATT | TTCGAGTCCA | AATGCCACAT | CGTATTTTAC | 1380 |
| | AATTGAACCA | ACAAATTGAT | TATCCGGATT | CTGAAATACA | ATTCCTATGT | CTTTTCTTAA | 1440 |
| 30 | CTTTTCAAAA | TTATCATCAG | TTATAGCTTG | ATTATTATAA | AAAATTTCTC | CAGATTTAAC | . 1500 |
| | TTTCTCTATG | CCAATCATTA | ACTTGGCAAT | TGTAGATTTT | CCAGAACCGT | TATGACCAAC | 1560 |
| | AATAGATGTC | CACTGACCTT | TAGGTATATT | AAAAGAAACA | TCTTTCAATG | TGAAGGATGC | 1620 |
| 35 | ATCACTTTGA | TATTGAAATG | AAACATTTTT | AAATACAATA | ACTGAATTCT | TATCCTCCAC | 1680 |
| | TIGICICICI | CCTTTACGAT | TCGTGTATCT | ATCATATTTT | ACAATATTTA | TAAATCGCTG | 1740 |
| 40 | TATATGACAT | TGACTGGGTT | CTCTATATAT | TACTAGTATT | TTCTGACTCA | TTTCTAGTCT | 1800 |
| 40 | TTAAAGTGTT | GTTTAACAAC | TAATGATAAG | GACTTTTATT | CCTCTCTAAC | AATTATGTAT | 1860 |
| | AAACGTTAAT | AAAATAAATG | ATTTACTAAT | ATAGGGGTGG | TCGCGTTTGA | TTCAACGATA | 1920 |
| 45 | ATACTTTCAC | TTCATTCAGT | TCTAGTGAAA | TTGATCAAAC | TAGCTTCATC | ATATTTTTAG | 1980 |
| | ATTCGCACTC | AAAAAAGTAA | ATATAAAGAA | ATCGGACTTA | AAAACATTTC | TGTTCATAAG | 2040 |
| | TCCGATATTT | TATTCAATAA | AAAAGCGCGC | ACCCCATCAT | AAGTTTGTTG | AGTTCACGCT | 2100 |
| 50 | TTAAATCTTT | ATTTAGTTGA | TGGGGTACTC | TGAGCTAGAC | AATATTTGTA | TGTGGCAAAC | 2160 |
| | ATTA + COTTO | CACTCATTTY | ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ | A ACTACTTAC | هريت لا متحت لا متتحت لا متتحت | TA A A TOTOTTA | 222 |

| | ACGAGTGTA | A CCACCTTGAC | GTTCTGTGTA | Acgetetaca | ATTTCACCAA | ATAATTTTTG | 234 |
|----|------------|---|------------|-----------------|-------------|------------|------|
| | AAGTGCAGTT | TGTGTAGTTT | CATCTTCGTT | TAAGATTTCA | ACATTACGTA | AAGTTTTAGC | 240 |
| 5 | TGCATTACGA | CGAGAAGCTA | AATCTCCTTT | TTTACCTAAA | GTGATTAATT | TCTCAACAAC | 246 |
| | ACTGCGAACT | TCTTTTGCAC | GAGCTTCTGT | AGTTTCAATA | CGTTCACTAA | TAATAAGTGA | 252 |
| 10 | TGTAGCTAAG | TCACGTAACA | TAGCTTTACG | TTGATCAGAA | GTACGACCTA | ATTTTCTGTA | 258 |
| 10 | ACCCATGAGT | TAACCTCCTT | TATCAATCTT | СТТТТСТТАА | TCCTAATCCT | AAATCTTCTA | 264 |
| | ATTTGTATTT | AACTTCTTCT | AAAGATTTAC | GACCTAAATT | ACGCACTTTC | ATCATGTCAG | 270 |
| 15 | CTTCAGATTT | GTCAGCTAAC | TCTTGAACAG | AATTGATTCC | TGCGCGTTTT | AAGCAGTTAT | 276 |
| | ATGAACGTAC | AGATAAGTCT | AATTCTTCAA | TAGACATTTC | TAATACTTTT | TCTTTTTGAT | 2826 |
| | CTTCTTCTTT | TTCAATCATG | ATTTCAGCGT | TTTGCGCTTC | ATCAGTAAGA | CCAACGAAGA | 2880 |
| 20 | TATTCAAGTG | TTCAGTCATT | ATTTTTGCTG | CTAATGAAAC | TGATTCTTGT | GGTGTGATTG | 2940 |
| | AACCATTAGT | CCAAACATCC | AATGTTAATT | TATCAAAATC | ACTGCTTTGA | CCTACACGTG | 3000 |
| | TATTTTCAAC | AGTATAGTTC | ACACGTTCAA | CAGGTGAATA | CAATGAATCA | ACAGGGATTA | 3060 |
| 25 | CACCAATTGG | TAAATCACTA | GTATTATTTT | GTTCTGCTAA | TGCGTAACCT | CTACCCTTGT | 3120 |
| | TAGCAACTAG | ACGAATTTTT | AAGTGACCAC | CTTTAGATAC | TGTTGCAATT | TTAAGCTCTG | 3180 |
| | GGTTTAAAAT | TTCAACATCA | CTATCATGTG | TAATGTCGCT | TGCTGTTACT | TCGCCTTCAT | 3240 |
| 30 | CACGTACATC | AATTTCTAAA | GTTTTATCTT | CTTCAGAGTA | AATTTTCAAT | GCTAATTGTT | 3300 |
| | TAATGTTCAT | AATAATTGTA | GAAACATCTT | CAACTACATT | GTCTACTGCT | GAGAATTCAT | 3360 |
| | GTAAAACTCC | CTCAATTTCA | ATATACTTAA | CGGCTGCACC | TGGTAATGAA | GATAGTAGGA | 3420 |
| 35 | TACGACGTAA | GGAGTTTCCT | AGTGTAGTAC | CGTAGCCACG | TTCTAGTGGT | TCAACAACGA | 3480 |
| | ACTÉACCGAA | TTTAGCATCT | TCACTAATTT | CAATTGTCTC | AATTCTAGGT | TTTTCGATTT | 3540 |
| 10 | CTATCATTTA | AATATCCTCC | TTATATACGT | CGACTTAATT | TAAAATGTTT | GCTCAGTGAC | 3600 |
| •0 | CTGTAACAAT | ACCATCATAA | ATTATACACG | ACGACGTTTT | GGTGGACGAC | AACCGTTATG | 3660 |
| | AGGTACTGGA | GTAACGTCTC | TGATCGCAGT | TACTTCTAAA | CCTGCAGATT | GTAATGCACG | 3720 |
| 15 | AATAGCTGAT | TCACGACCTG | GACCAGGTCC | TTTAACTGTT | ACTTCAACTG | TTTTTAAACC | 3780 |
| | ATGCTCCATA | GCTGATTTAG | ATGCAGTTTC | AGAAGCCATT | TGTGCTGCAA | ATGGTGTTGA | 3840 |
| | TTTTTTAGAT | CCTTTGAATC | CTAATGCACC | AGCTGATGAC | CATGATAAAG | CATTACCGAA | 3900 |
| 50 | CTCATCAGTG | ATAGTTACAA | TAGTGTTGTT | GAATGTTGAA | CGGATGTGTG | CTACACCATT | 3960 |
| | TTCAATATTC | TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT | TACCTTTACC | ACATIA CTVTVOTI | mma comocos | | |

| | CGCGCGTkGT | TTTTCGTTTT | TTGACCACGA | ACTGGTAAAC | CACGACGGTG | ACGGATACCC | 4140 |
|----|---------------|------------|---------------|--------------------|------------|------------|-------|
| | ACGGTATGAT | GAAATTnCCA | TTAAACGTTT | GATATTTAAG | TTAGTTTCAC | GACGTAAGTC | 4200 |
| 5 | ACCTTCGACT | TTATAACCGT | CTACAACTTC | ACGGATGCGA | CCTAATTCGT | CATCAGTTAA | 4260 |
| | ATCTTTCACA | CGAGTATCAG | CTGATACGTT | AGCTTCTTCA | AGAATTTTTT | GAGCAGTTGA | 4320 |
| | CGTACCGATA | CCGTATATAT | AAGTTAATGA | GATAACTACG | CGTTTTTCAC | GTGGAATATC | 4380 |
| 0 | TACTCCTGCA | ATACGTGCCA | TATTAATTTA | CACCTCTCTT | TTATTAACCT | TGTCTTTGTT | 4440 |
| | TGTGTTTTGG | ATTTTCACAA | ATTACCATTA | CTTTACCTTT | ACGTTTAATG | ACTITACATT | 4500 |
| 5 | TTTCGCAAAT | AGGTTTTACT | GATGGTCTTA | CTTTCATTTT | TATACCTCCC | TATATTATGG | 4560 |
| | AGTGACGATT | ATTTATAACG | ATAAGTAATT | CTTCCGCGTG | TTAAATCGTA | CGGAGACATC | 4620 |
| | TCAACAGTTA | CTTTGTCGCC | AGGTAGAATA | CGAATGTAAT | TCATTCTGAT | TTTACCACTT | 4680 |
| ?0 | ACGTGAGCnA | AAATCTCATG | ACCATTTTCT | AATTCTACTT | TAAACATTGC | GTTCGGTAAA | 4740 |
| | GTATCTAATA | CAGTACCTTC | TAATTCAATT | ACATCTTGTT | TAGCCATTGA | TTAACTTCCC | 48,00 |
| | CCTTTTTGCA | ATAGTAAGGT | AATCGTCAAT | AGACAACTTT | ATTGTTACGA | ATCTATCAGT | 4860 |
| ?5 | GATTAATTTT | ATAAGTTAAA | CAAAAATTAC | GGGAATTAAT | TATCGTTAAT | TGCCACTCTC | 4920 |
| | ATCTATCTAA | TATGATTAAA | TCATGCCTCA | CTTAAAATAG | ACCGCTAAAA | GTTGATCTAT | 4980 |
| 30 | TACAAATGAT | CTAAAATATC | AATGACATCT | TTGGTAACGT | CGCTAATATC | TTTTGAACCA | 5040 |
| | TCAATATTTT | TCAATACACC | TTTTTGATCA | TAGAAATCTA | AAATAGGCTT | AGATTGTTTA | 5100 |
| | ATATTAACAC | TCAAACGATT | AGCTACCGTT | TCAGGATTAT | CATCTTCTCG | TTGATACAAT | 5160 |
| 35 | TTACCACCAT | CGATATCACA | AATACCTTCG | ACTTCGGAGG | ATTAAATACA | AGATGATACG | 5220 |
| | TTGTACCACA | TGACTCACAG | ATTCGACGAC | CTGTAAGACG | GTTCATTAAT | TCTTCTTCCG | 5280 |
| | GAACTTCGAT | ATTGATGACA | GCATCAATGT | TTCTGTCAAG | CTCAGACATA | ATATTATTTA | 5340 |
| 10 | ATGCCTCAGC | TTGCTCGATT | GTTCTTGGGA | AGCCATCTAA | TAAAAAGCCT | TTTTTTGCAT | 5400 |
| | CGTCTTCAGA | AATTCTTTCC | TTAACGATAC | CTACAGTCAC | TTCATCAGGA | ACTAATTCGC | 5460 |
| | CACGGTCCAT | ATAAGACTTA | GCTTCTTTAC | CTAATTCAGT | TTCTTCTTTT | ATAGCTTTTC | 552 |
| 45 | TGAACATGTC | ACCAGTTGAA | ATGTGGGGTA | TTGGGAATTT | CTTGaCAATT | TCACTTGCTT | 5580 |
| | GAGTTCCTTT | ACCTGCGCCA | GGTAAACCCA | TCAAAATGAT | ATTCATAAGT | GCCCTCCTAA | 5640 |
| 50 | AATTATCTAC | CACCAAAGCC | TTTATATTCT | TTTTGAGATA | CTTGAGCTTC | TAAAGATTTC | 570 |
| | ATTGTTTCAA | TCGCTACACC | AATAACGATA | AGT AAACTTG | TACCACCAAT | CTGAATTGAT | 576 |
| | TOTO OTA A TO | CCATAAACTT | ACTOR OF A AT | ATCCCTACAA | TTCAAATAAC | CCCTNACNAC | 502 |

| | CCAGGTCTAA | TACCTGGAAC | ATAGCTACCT | TGTTTCTTAA | GGTTATCAGC | CATTTTTTCC | 5940 |
|---------------|------------|-------------|------------|------------|------------|------------|------|
| | GGATTAACTT | GTACAAATGC. | ATAGAAGTAT | GTGAATAGTA | TAATTAGTAC | AATATATACA | 6000 |
| 5 | ACCATACCAA | CATTACTTGA | AGGATTTGCA | GCATTCGCAA | TGTTTTGTGC | CCATTCTTTA | 6060 |
| | TCTGGATAGA | ACAACGTTAA | TGTTCTAGGC | AGTAAGAAGA | ACGCCATTGC | AAAGATTACA | 6120 |
| | GGAATAACAC | CGGCTGAGTT | CACTTTTAAA | GGTAGATAAG | TTGCCTGTGA | ACCTAATCTT | 6180 |
| 10 | TGAGCAGTTT | GTTTCTTAGC | ATATTGAATC | GGAATTTTAC | GAACGGCTTC | AAGTACATAA | 6240 |
| | ATAGCACCTA | CAGTTAATAG | TATCAGTGAC | ACTAAAAGTC | CTAATACTTT | CAACCATGCT | 6300 |
| 15 | AATGATGTAT | CTTCTTGCCC | AACGAACGCA | tTTGTcCAAA | TTGAATTAGA | CTGGCTGGCA | 6360 |
| | ACGTTGATAA | AATACCCGCA | AATATGATAA | TAGAAATACC | ATTACCAACA | CCGAACTGAG | 6420 |
| - | TGATTTGATC | ACCAAGCCAT | ATTAAGAAAG | CAGTTCCTGC | TGTnCAAAAC | TAGTGCTATT | 6480 |
| 20 | AATAAATAAC | TCATAATTGA | CTGATTGATA | ATCAGCGCAC | CTTTGAGATA | ATTATTAAAT | 6540 |
| | TGGAATGCCA | TACCTATAGA | TTGGATAAAT | GCTAAAGAAA | TTGCTAAATA | ACGAGTAACG | 6600 |
| | TTATTTAACT | TTCTTCTACC | TACTTCACCT | TGTTTTGCCC | ATTCTGAGAA | TTTAGGGACA | 6660 |
| 25 | ATATCCATTT | GTAATAATTG | CATTACGATT | GATGCAGTGA | TGTAGGGTAC | AATACCCATT | 6720 |
| | GCAAAAATAG | AAAATCGTTT | CAAGGCTCCG | CCACCAAAAG | TATTTAATAA | CTCAGTGGCA | 6780 |
| 20 | CCTTGAGAAC | CTTGGGGATT | ATCAAAAGCT | GCAGGATTTA | CTCCTGGAGC | TGGTATATAA | 6840 |
| 30 | GTCCCTATTT | TAAAAATTAC | TAACATTGCT | AGTGTGAAGA | AAATCTTGTT | ACGAACCTCT | 6900 |
| | TTTGTTCTAA | AGAAGTTCAC | aagggtttga | ATCATTAGAT | CACCTCGTGT | GCTCCACCTT | 6960 |
| 35 | TAGCATCAAT | AGCTTCTGCT | GCTGAAGCTG | AGAATTTATG | AGCTTTCACT | GTCAATTTCT | 7020 |
| | TATCAAGTGA | ACCATTACCT | AGTATTTTGA | TACCAGATTT | TTCATTCTTA | ACAACACCAG | 7080 |
| | ATTCTACTAA | TAAAGCTGGA | GTTACTTCAG | TACCATCTTC | AAATTTATTA | AGTTGGTCTA | 7140 |
| 40 | AGTTAACAAT | AGCATATTCT | TTACGATTTA | TGTTAGTAAA | ACCACGTTTT | GGTAAACGAC | 7200 |
| | GGAATAATGG | TAATTGACCA | CCTTCAAATC | CTGGTCTTAC | ACCACCGCCT | GAACGAGCTT | 7260 |
| | TTTGACCTTT | GTGTCCGCGA | CCACTTGTTT | TACCGTTACC | TGTCGCAACA | CCACGTCCAA | 7320 |
| 45 | CACGATTGCG | TTCTTTACGT | GAACCTTCTG | CCGGTTTTAA | CTCATGTAAT | TTCATTTCGG | 7380 |
| | CACCTCCTTG | ATTATTTTTC | TTCTACTGTT | ACTAAGTGCT | TAACTTTGTT | GATTTGCCCA | 7440 |
| 5 <i>0</i> | CGAATAGCAG | GGTTATCTTC | AACAACTACT | GAACTGTTAG | TCTTTTTAAG | ACCTAAAGCT | 7500 |
| 90 | TCAACAGTTT | TACGTTGTGT | TTCAGGACGA | CCAATAACAC | TACGAGTGAG | GGTAATTTGT | 7560 |
| | AATTTAGCCA | TAACTAGTTT | TCCCTCCTTA | ATTGTATAAT | TCTTCTACTG | TTTTGCCACG | 7620 |

| | CATGTTGATT | GGTGTGTTTG | ATCCTAATGA | TTTACTTAAG | ATATCAGTGA | TACCTGCTAA | 7740 |
|-----------|------------|------------|------------|------------|------------|---------------|------|
| | TTCAAGTACG | GCACGAACAG | GACCACCAGC | GATAACTCCT | GTACCAGGTG | CAGCCGGTTT | 7800 |
| 5 | CATAAATACG | CTTCCTGAAC | CGTAACGGCC | AGTAATTGTG | TGTGGAGTTG | TACCTTCAAC | 7860 |
| | ACGTGGAACA | ACTACTAAAT | CTTTTTTAGC | TGCTTCAACA | GCTTTTTTGA | TTGCTTCTGG | 7920 |
| | TACCTCTTGA | GCTTTACCAG | TACCGAAACC | TACACGACCA | TTTTTGTCTC | CAACTACAAC | 7980 |
| 10 | TAATGCAGTG | AAACGGAAAC | GACGACCACC | TTTTACAACT | TTTGCTACAC | GGTTGATTGT | 8040 |
| | AACAACGCGT | TCTTCAAATT | CTTTCGTCTC | TTCTTCtCTA | CGAGCCATGT | ATTTGTCCCT | 8100 |
| 15 | CCTTTAAATT | AAAATTCTAA | TCCGCTTTCT | CTTGCTGCTT | CAGCTAATGC | TTTAAcACGT | 8160 |
| 73 | CCGTGATATA | AATATCCTCC | ACGGTCAAAT | ACGATTTCTT | TAATGCCTTT | GTCAGCAGCT | 8220 |
| | TTTTTAGCAA | TTGCTTCACC | GACTTTAGTT | GCTAATTCAA | CTTTAGTTGC | TGTAGTAGCA | 8280 |
| 20 | ATGTCGCTGT | CTTTTGAAGA | AGCTTGAGCT | AATGTTACGC | CTTTATTATC | ATCAATAATT | 8340 |
| | TGAGCGTAGA | TATGCTTGTT | TGAACGATAT | ACGTTTAAAC | GTGGCTTTTC | AGCTGTACCT | 8400 |
| | GATAAGTTAG | TACGAACACG | AGCATGTCTT | TTTAAACGCA | CTTTATTTTT | ATCAATTTTA | 8460 |
| 25 | CTGATCATTT | CAATACTCCT | TTCTTTAGAG | TTTATCTATT | ATTTACCAGT | TTTACCTTCT | 8520 |
| | TTACGGCGAA | CGTATTCACC | TTGGTAACGA | ATACCTTTAC | CTTTGTAAGG | CTCTGGAGGT | 8580 |
| | CTTACTGAAC | GGATGTTAGA | TGCTAATGCT | CCAACTTGTT | CTTTTGAAAT | ACCTTCAACT | 8640 |
| 30 | TTAACGACTG | TGTTTTTCTC | AACTGAGAAA | GTAATGTTTT | CTTCAGCTTT | AATTTCTACT | 8700 |
| | GGGTGAGAAT | AACCAACGTT | AAGGATTAAG | TCTTTACCTT | GCATTTGAGC | ACGGTAACCT | 8760 |
| 35 | ACACCAACAA | GTTCAAGTAC | TTTTACGTAT | CCTTGAGAAA | CACCTTGTAC | CATATTGTTT | 8820 |
| 55 | AATAAAGCAC | GAGTTGTACC | ATGGTTTGTT | CTATCTTCTT | TAGAATCAGA | TGGTCTTACA | 8880 |
| | acttéaattg | TGTTTTCTTC | TTGTTTGAAT | GTCATTCTTT | CATTTAAAGT | TCTTGATAAT | 8940 |
| 40 | TCACCTTTAG | GACCTTTAAC | AGTTACATGA | TTTCCATCAA | AAGTTACTGT | TACGTCACTA | 9000 |
| | GGGATGTCAA | TAATTTTCTT | ACCAACACGA | CTCATGTTAT | GGCACCTCCT | TATTTTTTAT | 9060 |
| | TACCAAACGT | ATGCGATAAT | TTCTCCACCA | ACATTACGTT | TTCTTGCTTC | TTTGTCAGTG | 9120 |
| 45 | ATTACACCTT | CAGAAGTTGA | TACTAATGCA | ATACCTAAAC | CATTTAATAC | TTTAGGCATT | 9180 |
| | TCGCTAGCTT | TTGCATAAAC | ACGTAAACCT | GGTTTTGAAA | TACGTTTTAA | TCCTGTGATA | 9240 |
| | ACACGCTCAT | CGTTTTGACC | ATATTTTAAG | AATAAACGAA | GTACACCTTG | TTTATCATCT | 9300 |
| 50 | TCTACGTATT | CAACATTTTT | AATGAAACCT | TCACTCTTTA | AGATTTCAGC | AATTTCTTTT | 9360 |
| | **** | ATTCACCTAA | THE REPORT | TOTTO | CCMTCTTTCC | CONTROL DE LA | 9420 |

| | TCTTTTTTAT | TACCAGCTAG | CTTTACGAAC | GCCAGGGATT | TGGCCTTTGT | AAGCTAATTC | 9540 |
|------------|------------|------------|---------------|-------------|------------|---------------|-------|
| | ACGGAAACAA | ATACGGCATA | ATTTAAATTT | ACGATATACA | GAATGTGGAC | GGCCACAACG | 9600 |
| 5 | TTCACAACGA | GTGTATTCAC | GAACTGCATA | TTTTTGTTTT | TTTTGTTGCT | TAGCAACCAT | 9660 |
| | TGAAGTTTTA | GCCACTTAAT | TAGCCTCCTT | TAAATAATTA | TTTACGGAAT | GGCATACCGA | 9720 |
| 10 | AGTTAGCTAA | CAATTCACGA | GCTTCTTCAT | CAGTGTTAGC | AGTCGTTACG | ATAACAATAT | 9780 |
| | CCATTCCTCT | AACTTTACTT | ACTITATCAT | AGTCGATTTC | TGGGAAAATT | AATTGTTCTT | 9840 |
| | TAACACCTAA | AGTGTAGTTA | CCGCGTCCGT | CAAATGCTTT | TTTAGAAACA | CCTTGGAAGT | 9900 |
| 15 | CACGTACACG | TGGTaATGAT | ACTGAAATTA | ATTTGTCTAA | GAATTCATAC | ATTCTTTCAC | 9960 |
| | CGCGAAGTGT | TACTTTCGCA | CCGATTGGCA | TACCTTCACG | TAAACGGAAA | GTCGCGATTG | 10020 |
| | aTTTTTTAGC | TTTAGTTACT | AATGGtTTTT | GACCAGTGAT | CAATTCTAAT | TCTTCAACAG | 10080 |
| 20 | CATTGTCTAA | TACTTTAGAA | TTTTGTACTG | CGTCACCTAC | ACCCATGTTC | ACAACGATTT | 10140 |
| | TATCTATTTT | TGGTACTTCC | ATTACTGAAC | TATAATTGAA | TTTTTTCATT | AAGTTTTCAG | 10200 |
| | TAACTTCAGT | GTTaAACTTT | TCtTTTaAAC | GGTTCaAAGT | GGGATCCTCC | TTTCaACTTG | 10260 |
| 25 | TtATTAATTA | TTAGAKTTAA | TTTCTTCGCC | AGATTTTTTA | GCGATACGAA | CTTTTTTACC | 10320 |
| | ATCAACAAAT | TTGTAACCTA | CACGAGTTGG | TTCGTTTGTT | TTAGGGTCCA | ATAATTGTAC | 10380 |
| 3 <i>0</i> | ATTAGAAACA | TGGATTGCTG | CCTCTGTTTC | TAAGATTCCA | CCTTCAGGAT | TTAATTGAGT | 10440 |
| | TGGTTTTTGG | TGTTTTTTCA | TAATGTTAAC | ACCTTCCACA | ACGACACGGT | CTTTTTTAGG | 10500 |
| | TAGAGTAGCA | ATTACTTTAC | CTTCTTTACC | TTTGTCTTTA | CCTGCGATAA | CTTTAACGTT | 10560 |
| 35 | GTCACCTTTT | TTGATATGCA | TGTGGGCACC | TCCTTATTTG | TATTGGTTGT | TATTAATTAA | 10620 |
| | AGTACTTCTG | GTGCTAATGA | TACGATTTTC | ATGAAGTTAC | CTTCACGTAA | TTCACGAGCA | 10680 |
| | ACAGGTCCGA | AGATACGAGT | ACCACGTGGG | CCTTTGTCAT | CACGGATGAT | AACACATGCA | 10740 |
| 10 | TTTTCATCAA | ATTTGATGTA | TGAACCGTCA | TTACGACGAA | CACCTGACTT | AGTACGTACG | 10800 |
| | ATTACAGCTT | TGACAACGTC | ACCTTTTTTA | ACAACGCCAC | CTGGTGTTGC | ATTTTTAACA | 10860 |
| | GTACATACGA | TAACATCGCC | GATGTTTGCT | GTTTTACGAC | CAGATCCACC | TAATACTTTG | 10920 |
| 15 | ATTGTAAGAA | CTTCACGAGC | ACCAGAGTTG | TCTGCTACTT | TCAAGCGTGT | TTCTTGTTGG | 10980 |
| | ATCATTAGTT | AAACCTCCCT | татстстааа | CTTGTATTAA | ATAATTACTG | ACTCTTCAAC | 11040 |
| - 0 | AATCTCTACT | AAACGAAAAC | GTTTTGTTGC | TGATAAAGGA | CGAGTTTCTT | GAATTTTAAC | 11100 |
| 50 | AATGTCTCCT | AATTTAGCTG | AATTGTTTTC | ATCATGAGTT | TTGTATTTTT | TAGAGTATTT | 11160 |
| | TACTCGTTTA | СССТВТВВТТ | TATACTATATATA | СТА ВСТТТСА | ACAAGTACTC | ጥል አጥል ርጥር ተም | 11224 |

| | TTTTGTAACC | TCCTCTTACT | TAATTATTGA | TTAGCCTTAC | TTTGTTCAAT | TTCTCTTTCA | 11340 |
|----|------------|------------|------------|------------|-------------------|------------|-------|
| | CGAGCAACAG | TTTTTAGACG | TGCAATCGTT | TTTCTTACTG | TACGAATACG | TGCAGTTTCT | 11400 |
| 5 | TCTAATTGAC | CTGTAGCTAA | CTGAAAGCGT | AGGTTAAAAA | GCTCTTCTTT | TGAAGATTTG | 11460 |
| | ATTTGTTCTT | CGATTTCTGA | AGTGGTTAAG | TCTCTAATTT | CCTTAGCTTT | CATTTGTTTC | 11520 |
| | ACCACCCAAT | TCCTCACGTT | TTACAAACTT | AGTTTTTACT | GGAAGTTTGT | GACTTGCTAA | 11580 |
| 10 | ACGTAGTGCT | TCACGCGCAA | CTTCTTCAGA | AACGCCAGCA | ACTTCGAATA | AAATTCTACC | 11640 |
| | TGGTTTAACA | ACTGCGATCC | AGCCTTCAAC | CGCACCTTTA | CCAGCACCCA | TACGTACTTC | 11700 |
| 15 | TAAAGGTTTT | TTAGTATATG | GTGTATGTGG | GAAGATTTTA | ATCCAAACTT | TCCCGCCACG | 11760 |
| | TTTCATGTAA | CGTGTCATTG | CTATACGAGC | AGATTCGATT | TGACGAGATG | TGATCCAAGA | 11820 |
| | CGTTGTTGTA | GCTTGTAAAC | CAAACTCACC | AAATGTTACG | TALTACCGCC | TTTAGAACGA | 11880 |
| 20 | CCAGTTGTTT | TAGGACGATG | TTGACGACGA | TATTTTACAC | GTTTTGGTAG | TAACATTATT | 11940 |
| | ATTTTCCTCC | TCCACTAGTG | TTCTTAGTAG | GAAGAACTTC | TCCACGATAA | ATCCATACTT | 12000 |
| | TAACGCCTAA | TTTACCGTAA | GTAGTGTCAG | CTTCAGCGTG | tGCATAATCG | ATGTCAGCAC | 12060 |
| 25 | GTAACGTATG | AAGTGGAACA | GTTCCTTCTG | AATATTGTTC | AGCACGAGCG | ATGTCAGCTC | 12120 |
| | CGCCTAAACG | ACCAGATACT | TGaGTTTTGA | TACCTTTAGC | ACCAAGTTTC | ATAGCTCTAG | 12180 |
| | TGATTGCTTG | TTTTTGTACA | CGACGGAATG | AAGCACGGTT | TTCTAATTGA | CGTGCGATGT | 12240 |
| 30 | TTTCAGCTAC | TAAACGAGCG | TCAAGATCAA | CTTTTTTGAT | TTCAATTACG | TTGATGTGTA | 12300 |
| | CTTTTTTATC | AGTTAACGCA | TTTAATTTGT | TGCGTAATTT | TTCGATTTCT | GAACCGCCTT | 12360 |
| 35 | TACCAATTAC | CATACCAGGT | TTACCAGTAT | GAATTGCAAT | GTTGATACGG | TTTGCAGCAC | 12420 |
| | GTTCAATCTC | TACGTGAGAA | ACTGATGCTT | CTTTTAATTC | ATTATCAATA | AATTTACGGA | 12480 |
| | TTTŢĀĀĀĀTC | TTCGTGTAAA | AGTGAAGCGA | AGTCTTTTTC | AGCATACCAT | TTAGCTTCCC | 12540 |
| 40 | AATCACGGAT | AATACCAACA | CGAAGTCCGA | TTGGATTAAT | TTTTTGACCC | ACAGTATTCC | 12600 |
| | CTCCTTAAAA | GTTAATTAAG | CTTCTTTAGC | TTCTTCTTTA | CCGTCACTTA | CGACGATTGT | 12660 |
| | AATGTGGCTT | GTACGTTTGT | TAATCGCACT | TGCACGACCT | TGCGCACGTG | GACGGAAACG | 12720 |
| 45 | TTTTAATGTT | GGTCCTTCGT | TAGCATATGC | TTCTTTAACT | ACTAATTCAT | CTGTGTTCAT | 12780 |
| | GTCATAGTTA | TGTTCAGCAT | TAGCTAAAGC | GGACATTAAT | ACTTTTTCAA | TTACTGGTGA | 12840 |
| | TGAAGCTTTG | TTTGTTAATT | TTAAAATTGC | AATAGCTTCA | GCAGCATTTT | TACCTCTGAT | 12900 |
| 50 | TAAGTCAAGA | ACTAGTCTTA | CTTTACGAGG | TGCGATTCTT | ATTGTTCTAG | CAACCGCTTT | 12960 |
| | TGCTTCCATT | AGGATGTCCT | CCTCTACTTA | ATAGATATTA | TCTTCTTGTT | TTCTTGTCGT | 13020 |

| | TATCTTCAGT | TACATATACA | GGTACGTGTT | TACGTCCGTC | GTATACTGCA | AAAGTATGTC | 13140 |
|------------|------------|-------------|------------|-------------|---------------|----------------------|-------|
| | CGATGAAATT | AGGGAAAATT | GTAGAACGAC | GTGACCATGT | TTTGATTACT | TGTTTCTTTT | 13200 |
| 5 | CGCTTCCTTC | TTGAGCTTCA | ACTTTTTCA | TTAAATGCTC | ATCGACGAAA | GGTCCTTTTT | 13260 |
| | TAATACTACG | AGCCATTTGG | GCGCCTCCCT | TCTTATTATG | TGCGTGCAGC | TTTAAGCCGC | 13320 |
| | ACACCCAAAT | AAGTTGATTA | TATTATTTTT | TCTTACGTCC | ACGAACGATA | AGTTTGTCTG | 13380 |
| 0 | ATGATTTTTT | ACCACGACGA | GTTTTCTTAC | CAAGCGTAGG | TTTACCCCAT | GGTGACATTG | 13440 |
| | GAGATGGTCT | ACCGATAGGA | GCACGACCTT | CACCACCACC | GTGTGGGTGA | TCGTTAGGGT | 13500 |
| 15 | TCATTACAGA | ACCACGAACT | GTTGGACGGA | TACCTTTCCA | TCTTGAACGT | CCGGCTTTAC | 13560 |
| | CAACGTTAAC | TAATTCGTGT | TGTAGGTTAC | CAACTTGACC | GATTGTAGCA | CGGCAAGTAG | 13620 |
| * | ATAAGATCAT | ACGAACTTCA | CCAGATCTTA | ATCTGATTAA | TACGTATTTA | CCTTCTTTAC | 13680 |
| 20 | CAAGTACTTG | AGCACTTGCA | CCAGCTGAAC | GAGCGATTTG | TCCACCTTTA | CCAGGTTTAA | 13740 |
| | GCTCGATGTT | GTGTACTACT | GTACCAACTG | GAATGTTTTG | TAATGGTAAT | GCGTTACCAA | 13800 |
| | CTTTGATGTC | AGCTTCAGCA | CCACTTTCAA | CGATTTGACC | TACTTCTAAT | CCTTTAGGAG | 13860 |
| ?5 | CAATGATATA | TCGTTTTTCA | CCGTCTGCAT | ATACAACTAA | AGCGATGTTT | GCTGAGCGGT | 13920 |
| | TTGGATCATA | TTGAATAGAA | TCAACTTTTG | CATTGATACC | ATCTTTGTTA | CGTTTGAAAT | 13980 |
| | CGATAACACG | GTATTGACGT | TTGTGTCCAC | CACCATGGTG | TCTTACAGTC | AATTTACCTT | 14040 |
| 30 | GGTTGTTACG | TCCCGCTTTT | TTCGGTAGCG | GTTTTAATAA | TGACTTTTCA | GGTGTAGTTT | 14100 |
| | TCGTGATTTC | TGCGAAATCT | AACGAaGTCA | TATTACGACG | ACCATTTGTT | ATTGGCTTAT | 14160 |
| 3 <i>5</i> | ACTITITAAT | AGCCATTGTC | GCTTACCTCC | TTAATGGTAA | TTGTTTTATT | AGTTAAATAA | 14220 |
| | GTCGATTGAT | CCTTCTTTAA | GAGTTACAAT | CGCTTTTCTT | CTTTTGTTTG | TATAGCCTTG | 14280 |
| | GTAACGGCCC | ATACGTTTTT | TCTTAGGTTT | GTAATTCATG | ATATTAACAC | TTGCAACTTT | 14340 |
| 40 | TACGTTGAAG | ATTTCTTCAA | CTGCCATTTT | TACTTGTGTT | TTGTTAACAC | GAGTATCAAC | 14400 |
| | GTCGAAAGTG | TATTTGTCTT | CAGCCATTGC | TTCAGAAGAT | TTCTCAGTGA | TTACGGGGCG | 14460 |
| | CTTAAGAATA | TCTCTTGCTT | CCATTATCCG | AGCACCTCCT | CAACTTTTTT | AGCAGCAGCT | 14520 |
| 45 | TCAGTAATTA | CTAAGCTGTC | AGCATTAGTG | ATATCTAAAA | CATTTAAACC | TTGAGCAGTT | 14580 |
| | GTCACTTGAA | CGCCAGGGAT | GTTGCGTGCT | GATAATTCAA | CATTTACATC | TTCGTTTTCA | 14640 |
| | GTAACTACTA | ATACTTTTTT | AGGTTGTTCT | AATGTAGATA | ATACATTTTT | GAATTCTTTA | 14700 |
| 50 | GTTTTTGGAG | CTTCGAAGTT | GAATGCGTCA | ACTACAGTTA | AGCCATTCTC | TTGAGCTTTG | 14760 |
| | ***** | CDC1 CCCD11 | | 00010000000 | m1.0001.00000 | CONTROL OF THE PARTY | 14000 |

| | CCTTGACGAG | CACGACCTGT | TCCTTTTTGC | TTCCATGGTT | TACGTCCGCC | ACCGCTTACT | 14940 |
|----|------------|------------|------------|------------|------------|------------|-------|
| 5 | GCTGAACGAT | TCTTAACAGC | ATGCGTACCT | TGACGTAATG | AAGCACGTTG | TAAATTAATA | 15000 |
| | GCTTCGAATA | AAACGCTATT | ATTTGGCTCA | ATACCGAATA | CTGCATCGCT | TAATTCGATT | 15060 |
| | GAACCTGATT | TAGTTCCGTC | TAATTTTAAA | ACATCATAAT | TAGCCATTAT | GCATTTCCTC | 15120 |
| 10 | CTTTCACTTC | TTATTATTTA | TTACCTTTTT | TAATTGAAGT | TCTGATTTCT | ACTAAACCTT | 15180 |
| | TTTTAGGTCC | AGGTACGTTA | CCTTTTACTA | AGATAACTTT | GTTTTCTGTG | TCAACTTGAA | 15240 |
| | CTACTTCTAA | GTTTTGAACA | GTTACAGTGT | TTCCACCCAT | ACGTCCTGGC | ATTITTTGGC | 15300 |
| 15 | CTTTAAATAC | TCTAGAAGCA | TCTGAAGCCA | TACCTACAGA | ACCTGGTGCT | CTGTGGAAAT | 15360 |
| | GAGAACCGTG | TGACATAGGT | CCACGAGATT | GTCCGTGGCG | TTTAATTGCA | CCTTGGAAAC | 15420 |
| | CTTTACCTTT | TGATACGCCT | GTTACGTCAA | TAACGTCGCC | AGCTACAAAA | GTATCTACTG | 15480 |
| 20 | AGACTTCTTG | Aacctactcg | TAAGCATCCA | CGTCTACATT | GCGGAATTCA | CGAATGAAGC | 15540 |
| | GCTTAGGTGC | TGCGTCAGCT | TTTTTAGCGT | GACCTTCAGC | TGGTTTATTA | GCATATTTAT | 15600 |
| | TAGATTTTGC | ATCTTTTTTG | TATGCTTTTT | TGTCTTCAAA | TCCAACTTGG | ATTGCGTTGT | 15660 |
| 25 | ATCCATCAAC | TTCTACAGTT | TTCTTTTGTA | ATACAACATT | TTCTTTAGCT | TCTACTACTG | 15720 |
| | TTACAGGGAT | TAATTCACCG | TTTTCTCCGA | ATACTTGTGT | CATCCCAATT | TTTCTTCCTA | 15780 |
| 30 | AGATTCCTTT | GGTCATCGAA | AGTCCACCTC | CTAAAATTGT | CTATTATAAT | TTGATTTCGA | 15840 |
| | TGTCTACACC | AGATGGTAAG | TTTAAGCCCA | TTAAAGCGTC | AACTGTTTTT | GGTGTTGGGT | 15900 |
| | TTACAATATC | GATTAAACGT | TTGTGTGTAC | GTTGTTCGAA | TTGTTCACGT | GAATCTTTAT | 15960 |
| 35 | ACTTATGCAC | GGCACGGATG | ATTGTGTAAA | CTGATTTCTC | AGTTGGTAAC | GGAATTGGTC | 16020 |
| | CAGAAACATC | TGCACCAGAA | CGTTTCGCTG | TTTCTACAAT | CTTCTCTGCT | GATTGATCAA | 16080 |
| 40 | TTACCCCGTG | ATCATAAGCT | TTTAATCTGA | TTCTGATTTT | TIGITITGCC | ATAATTTTCC | 16140 |
| | CTCCTTATTC | GTCTACATTT | AGTGATAGAC | TTCTCCACGA | AAACTATCTT | ACACAGCGCC | 16200 |
| | ATGGCAAAGC | GGCCGGGTGT | GTCAGTAACC | TTTCGCTTCA | TCGCTTTTCT | TAAAGTCCAA | 16260 |
| 45 | CGTTAGTTAT | ATTACACGAA | AAACATCGAT | AAATCAAGGC | TTTTCACATA | ATTITICIAT | 16320 |
| | CTGTCTAACA | CATACTTTTA | TATTTDACTT | TATATACTTA | GTCAGTTCAA | CTATTTTCGA | 16380 |
| | GATATTTTNA | ATTTCCn | | | | | 16397 |

(2) INFORMATION FOR SEQ ID NO: 206:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29555 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 206:

| J | Thagttgttt | CTGCCACGAA | AGATTCAATG | GCTTTTCTTG | CTTTACGCTT | TTCTTTCAAT | 60 |
|----|------------|------------|-------------------|------------|-------------|------------|------|
| 10 | GGCAAATCAC | CAATCATTTT | TTTAAGGTGA | TGTGGGTTTA | CAACACCACT | ATACTGGTAG | 120 |
| | TCATTTGAAn | TTGTTTTTAG | GGCTTGTTCA | TCGATAGATC | TCTCTCCAGC | AAATCCTTTG | 180 |
| | AACTCCGCTT | CTTTTTTAAT | ACTTTCGAAA | TTAACATATT | CTTGATCGAT | ATCATCATCC | 240 |
| | TTATTTAAAG | AAGGTACAAC | ATTGTCGATG | AATTCTCTAA | TTAGATCTCG | TTTTAACCTC | 300 |
| | AATGLCGGAT | CATCTGCATG | ATCTAAAATG | CGTCTAATTT | GTTCTTGGTT | ACGACGTTGT | 360 |
| | TCCGCTTTGT | CTTCAAGATC | AATTTGTCTC | AATATATTCA | TAATATAATT | CACATTAATC | 420 |
| 20 | GTATCATTAC | GCATCATTTC | TATTTCGAAA | TCAATATCAT | TTAAAATGGA | TACTTTATTT | 480 |
| | TTCTCAGCCG | TCGCTCTTTT | TACTTGATCG | TACACAGCTA | AATATTTACT | TTTATAGTCT | 540 |
| | TCATTCTCTT | GTTCATCCAT | TCCAATTTCA | TCAATTGTAA | ACTCAAACTC | GTCAAATGCT | 600 |
| | TTTAAACGTA | ATATTATTTT | AGCTAATAAA | CGATAAGCET | CAACAAAGCG | CTTTAGCTCT | 660 |
| 25 | TCTTCATCET | GaATGtCATC | AACCATGTGT | GGTGTCGGCA | CAATCATTTT | AAGCTCACGA | 720 |
| | TAAGCGTCCA | TAAATTCTTT | TTTATACTCT | TCATAACTGC | GCATTAAAAT | TGTATCCGTA | 78.0 |
| 30 | TCATTTGTTT | GTGAGAATAC | TCTCAGTGCA | TCGTCTGTCT | CTTTTTTCAA | GTCACGATAG | 840 |
| | TTTACAATTT | TACCAAATGG | CTTTGATTCT | TTTTCAACCC | TATTTGTACG | TGAATACGCT | 900 |
| | TGAATTAAAT | CATGATACAT | TAAATTCTTA | TCAACATATA | AAGTGTTCAG | TACTTTACTA | 960 |
| 35 | TCAAAACCAG | TTAAGAACAT | ATTAACAACG | ATTAAGATAT | CAATTTTACT | ATCTTTAACG | 1020 |
| | CCCTTTTTAA | CGTTTTTtGA | AATATGATTA | AAATACTCAT | TAGTTGtGGC | TGnTGaAAAA | 1080 |
| | TTCGTCTCGA | ACTITITATT | ATAATCACTA | ATCATTATCT | CTAATTTTTC | ACGTGAATGA | 1140 |
| 40 | TATGGCACTT | CACCATCACG | ATCATCTTCA | TTAGGTTTAA | ACGTAAATAT | ACCAGCTATC | 1200 |
| | GTTAACGGTT | GTTCCAACTT | TTTGTTAAGT | CGCTTAAATG | TCTCATAATA | TTTAATAAGC | 1260 |
| | GCGTGAATAC | TTTGGACTGT | AAATATACTT | GAATATTGAC | GATTACGTGT | ATATTTATCA | 1320 |
| 45 | TGATTATTGA | TGATATGTCG | TGTTACTAAT | TCCACACGTT | TATCCGCTAA | CCATACTTCT | 1380 |
| | TCCGTATCAA | TTGCTTCAAC | CATGCtGTTA | TCTTCTGCTT | TTAAAGCTTT | AAAATTTTTA | 1440 |
| 50 | GTATTAATAT | AGTCAACTGA | GAAACCAAGT | ACATTACCAT | CATGAATGGC | ATCTCTAATT | 1500 |
| | AAATACGTAT | GTAAGCATCT | ACCGAAAATA | TCTGCAGTTG | TTCTACCATC | TTGACTACTA | 1560 |
| | THETCHGGAA | AACCTCCCCT | АССАСТСААТ | ССАВАСТАТТ | GGGCATTTTTT | GAAATCTTCT | 1620 |

| | ACTITATTCG | TTTTATACTG | TTCTAATAAA | GGGGCATTCC | CTTGAATCGC | TTTAGCCATT | 1740 |
|----|------------|------------|------------|------------|------------|------------|------|
| | TTTTGAATCG | TCGTTACAAT | AAGTGGCAAA | CTTTTATCAT | TTAGTTGGCG | TACCAGTTGC | 1800 |
| 5 | GAGGTATTAA | AAGTTTTGTC | TACAGCACCC | TTAGCAAATT | TATTAAATTC | CTCTTCTGTT | 1860 |
| | TGACTATCCA | AGTCTTTACG | GTCAACCAAA | AAGATAACTT | TCTTAATGTC | ATCTTGCTGT | 1920 |
| | GATAAAATCT | GACTCGCTTT | AAAAGAAGTC | AACGTCTTAC | CACTTCCAGT | TGTATGCCAT | 1980 |
| 10 | ACATATCCAT | TATTCCCTGT | CTCAGTCGCT | TGTTGAATAA | GTGCTTCTAC | CGCATACACT | 2040 |
| | TGATACGGAC | GCATTGCCAT | CAGTATTCTA | TCTGTTTCAT | TAATAATCAT | ATAGCGCGAT | 2100 |
| 15 | ATCATCTTAG | CTAATTGACA | AGGTCTCATA | AATGACTCAG | CAAACGATTG | CAATGTATTG | 2160 |
| | ATACGGTTAT | TCTGTTTATC | ACTCCAATAA | AACATGTGAC | TCTTCAATAG | TTCGCTATCA | 2220 |
| | TTATTAGAAA | AGTATCGCGT | TTCAACACCA | TTACTAATGA | TAAACATTTG | TATGTAGCGG | 2280 |
| 20 | AATAAGCCTG | TGTAATTTTG | TTTGCGGTAA | CGTTTTACTT | GGTTAAACGC | CTCATTAATA | 2340 |
| | TCAATACCTC | GACGTTTCAA | TTCAACTTGG | ACAAGGGGTA | GTCCGTTGAT | TAATATCGTT | 2400 |
| | ACATCATAAC | GTGCTTTATA | TGTATCCTCG | ACAGATACTT | GATTCGTCAC | TTGAAACTTA | 2460 |
| 25 | TTTTTACACC | AACTTTTCGT | ATCTAAAAAC | GACAAATAAA | TCTCAGACTC | ATCATCACGT | 2520 |
| | CTAAGTGGTA | ATTTATCACG | TAAAATACGG | GCACTCTCGA | AAATACTTTT | TCCATCAATC | 2580 |
| | ATCGTTAACA | GACGTTGaAA | TTCTTTATCT | GTTAaGGGAT | TGCCTTCTAA | TTTGTCCGCA | 2640 |
| 30 | TGACGCTCAT | TTAAAATCGT | TCTAAAATTA | TCAAGCAATT | GCTTATTATC | ACGTATCGTT | 2700 |
| | ACTCTTTCGT | AACCCAATTG | TTCAAGTTGA | TTCATCATTT | CATTTTCTAA | TGCGTATTCA | 2760 |
| 35 | CTTTGGTATG | CCATTCATAT | CCCCTTCCAT | ACACTTTCTA | TTGCTCTAAA | TATATCATAA | 2820 |
| | ACTTTAATGA | AAAATGTTTG | TTTTTTATCT | TCAAACGTAA | ATTTATTCTA | ATTTTATTGT | 2880 |
| | CTTATCTTTT | AATATTTGTC | TTTGAGGTAA | GTCGTATACT | AAAATTTGAA | TACAAATAAT | 2940 |
| 40 | CAAATCATTG | ATAAATTTTT | TGTCTACGAT | TAATGGAGGG | ACTTGAATGG | TGTTAATTAC | 3000 |
| | CTATCAAATC | ATTTTATTTT | TTATTATTAG | TCTAAGTTAC | TATTTAACTT | TAAATCATTA | 3060 |
| | CATGGCAGTC | ACTGTAGGTA | ACTTCACTTC | AATATTCGGC | ATGTTCGCAG | CCATACTCTT | 3120 |
| 45 | TATGTACTAC | TACCTACTCT | ATAAAAGTCC | CGAATACAAT | CAACGCAAAC | GATTTAAACA | 3180 |
| | TTTCATTCAT | ATCACTAATT | TGATAATAAT | TGCTTTTAGC | ACCTTCGTAT | TAGTTCATTT | 3240 |
| | AGCATTAAAA | TTATTCTTCA | GCATTTAATT | TCCATCTATG | AAAAAAGCAA | AGCTCAAATC | 3300 |
| 50 | TGAACTTTGC | TTTAATTTGT | CACGCCTTTA | TCATTTTCAA | AATAGCCTCT | ATGCCAGTTT | 3360 |
| | тасааастту | TAGCAACAAT | TTTTCATCAA | CCAACTGAAT | CACATCAAAA | ACTTCAATTG | 342 |

· 55

| | GICGCAAGAI | GCTTCCTGTA | ATTATCAAGT | GCCATTTTCG | ATTGGGTTAT | ACAATCTAGA | 354 |
|------------|--------------|------------|------------|------------|------------|------------|------|
| | . ATCGCATGAT | AATTTAATGC | TACAAATCGA | TAGTACAATA | TATCTACCGT | GAATAACTGT | 360 |
| 5 | GCAAATAGTG | ACGTTGTAGC | CGCCATACGC | ATTTCATTTT | CATCAGTTCT | GCCATAAATC | 366 |
| | AATGCATAGT | CTGCAATTTG | AGCCACTGGA | TTATTAGCTG | TACTAGATAT | AGTTATGATG | 372 |
| 10 | GGAATACTGT | AATGTGTGGC | CACCTGTGCA | ATTGACTGCA | ATTCACTATG | ACTACCTTGA | 378 |
| ,, | TTCGTCACAA | AAATCATGCA | ATCTCTATCA | TCATGCGTCG | CAAATGTTGA | CACAAGTAAA | 384 |
| | TGCGTTTCAT | GTAATAACCT | GACATTTAAG | CCAATACGAG | ATAACTTTTG | AAAAAGATCA | 3900 |
| 15 | CCAATAGTCA | AACTCGATGC | GCCAAATCCA | AATAAAAATA | TTGTCCTGGc | ATTTTTCAAC | 3960 |
| | ACATCACAAA | TTGCATCAAT | TTGCGCATCC | ATAATATTAG | TAGCTACAAA | TCGCATCGTA | 4020 |
| | TTCGTTGCTC | TAGCAATCAT | TTTATTTTTC | AAAGTTTCTA | CAGATTCATT | TTCAATCAAT | 4080 |
| 20 | TCTAAATGTG | GATTGGTTGC | AATATCTTCG | GGTAAGTATC | GAGATATCGC | AATCTTTAGC | 4140 |
| | TCTTGAAAAC | CTTGATGTGT | CATTTTCCGA | CTAAATCTAA | CAATTGATGC | TGTACTAACA | 4200 |
| | TTCGTAACAT | CTGCCAAATC | ATTCACAGTC | ATATCAATGA | TTTTATGTGG | ATTCTTTAAA | 4260 |
| 25 | ATGTAATCAG | CGATTATCTT | TTCTGTCTTC | GTAAAATCAC | TCAACTGCTT | ATCAATGCGA | 4320 |
| | TATAAAATAT | TTGTCATCAT | TAATCACCCA | ACAAATCTGT | CTGTCGCATC | GCCTTTGTCG | 4380 |
| 30 | TTCCAAATAA | ATATGTACAA | ACGAATCCAC | CAGCATACGC | AGCAAGTAAT | CCTGCAATAT | 4440 |
| | AACCTAAATA | CATATTATCT | GAGATTAATG | GTAATAGTGA | CACACCACTT | GGGCCTATTG | 4500 |
| | CTTTGGCACC | AATATGTCCA | ATTCCACCTA | TTACAGCGCC | ACCAATACCA | CCACCAATAC | 4560 |
| 35 | AAGCAGTTAA | GAAAGGTCGA | CCTAATGGCA | AAGTCACACC | ATAGATTAAT | GGTTCTCCGA | 4620 |
| | TACCTAGGAA | ACCAACTGGC | AATGCACCTT | TTAAAGTATT | ACGTAATGTT | GTGTTGCGTT | 4680 |
| | TACATCTTAC | CCAAAGTGCT | AATGCGGCAC | CTACTTGTCC | AGCACCAGCC | ATCGCTGCAA | 4740 |
| 10 | TTGGCAATAA | GTAAGTAGCA | CCTGATTGGT | TAATCATTTC | TATATGAATT | GGCGTAAAAA | 4800 |
| | TATGATGAAG | CCCTAACATA | ACTAACGGTA | GGAAGCTTGC | ACCAATGATA | AATCCACTAA | 4860 |
| | ATACGCCACC | AATACTAATA | ATTCCGTTAA | CTACTGAAAC | TAAACTGTCT | GAAACAAAAC | 4920 |
| 15 | CTGCTAATGG | CATAAAGATA | AAGATAGTTA | ATAGTCCTAC | AATCAACAAT | GCAATAGTCG | 4980 |
| | GCGTTACAAT | AATATCAATC | GCATTTGGCA | CAATTTTATG | TAATCTCTTT | TCGACAATAC | 5040 |
| 5 <i>0</i> | TTAAAATCCA | AACGGCAAAA | ATAACGCCAA | TAATCCCACC | TTGTCCAGGT | TGCAATGGTT | 5100 |
| | CTCCAGTGAA | GACATTCATT | AAAATATTTT | TACCAGCAAT | ACCCGTTAAT | AACGTTGTAC | 5160 |
| | CACCAATCAC | GCCACCAAGT | CCTGGTGTCG | CACCAAATTC | TTTAGCCGCA | TTAATACCAG | 5220 |

| | GCGTAATCCA | AGCACCTGAA | ATATAGCCTG | CCACCATTAA | GTTACTCAGT | ACTGCTGCAA | 5340 |
|----|-------------------|--------------|------------|--|------------|------------|--------------|
| | TACCACCAAT | TAATCCAGCT | CCAATAAATG | CAGGAATCAA | CGGTATAAAG | ATATTGGCAA | 5400 |
| 5 | TTGATTTCAA | TACTTTATTC | AACTTACCAT | TCTTTTGTTT | TGCTTTATGC | GCTTCCTTAT | 5460 |
| | TCGCCTTTGC | TTTATCAGCT | GCATATGATT | TATAGTCCAT | TTTTTCACTA | TCATTGTGAT | 5520 |
| | GGTGTGGTAT | TGGGTCACCT | AGTTTAACAC | CACTTAATTC | CGCCATATGA | TTAGCCACTT | 5580 |
| 10 | TATTGatGTA | CCAGGTCCAA | CCACAACTTG | AATGCGTTCA | TCGTGTATAA | CACCCATGAC | 5640 |
| | ACCATCAATA | TGCCTTAGTT | CTTGGTCATC | TACTTTATTC | TCATCTAATA | CTTTAATACG | 5700 |
| 15 | CACACGTGTC | ATACAGTTCA | TGACACTATC | TATATTATCC | ATACCACCTA | CTGCAGCAAT | 5760 |
| | AATTCGTTCT | GCAAGTTGTT | GTTCTTTGGT | CATTTAAATC | CCTCCTAAGG | TTGTCTATCT | 5820 |
| | CTGATTGCTC | GTTTAAaATG | TCACCATTGT | TTAATAACCG | TCTTGTTGCT | TCTTCCTTAG | 5880 |
| 20 | AAATGCCACA | CATACCCATA | ACTGTCGCAA | CTTTCACATC | ATGCTCAGAT | ACCTGATATA | 5940 |
| | ACGCCATTGC | TTCATCATAT | GTGATAGCAC | ATATTTCTTG | AATAATACGC | ACTGAACGGT | 60 00 |
| | CGATCAGTTT | TTGATTGGTT | GCTTTAACAT | CAATCATGAG | GTTATCGTAA | ACTTTTCCGA | 6060 |
| ?5 | CACCAACCAT | TGTGATGGTT | GAAATCATAT | TTAAAATTAA | CTTTTGTGCt | GTACCAGACT | 6120 |
| | TTAAACGTGT | TGAACCAGTT | AATACTTCTG | GACCAACTTT | AACTTCTACT | GGATACTGCG | 6180 |
| | CAATTTCACT | TATAACTGCA | TGTTCATTGC | ATGAAATAGA | TACTGTTGTA | GCACCGATTG | 6240 |
| 30 | TGTTAGCAAA | TGTTAAACCG | CCTATAACAT | ATGGCGTTTT | GCCACTCGCG | GCAATTCCTA | 6300 |
| | TAACGACATC | TTTTGATGTT | AAATCTATAT | TTTTCAAATC | TTCTTCCGCT | AATTTTTTGT | 6360 |
| 35 | GATCTTCCGC | ACCTTCTACA | GCCATCGTCA | TAGCATGTTG | TCCACCAGCA | ATAATACCTA | 6420 |
| | TAATTTCATG | AGGGTCAGTA | TTGAATGTAG | GTACACACTC | CGCTGCATCT | AAGACACCCA | 6480 |
| | ACCTICCACT | TGTACCTGCA | CCGATATAAA | TCAATCGTCC | ACCCTTTTTA | TACTGTGCAA | 6540 |
| 10 | TIGITITITI | AATTACTTTT | GTCAATTGTG | GTATTGCCTT | TCGAACTGCT | AACGGGACTT | 6600 |
| | GCTGATCTTC | TTTATTCATC | GTAATTAAAG | CCTCTTCCAC | AGTCATTTCA | TCAAGATGCA | 6660 |
| | TCGTCGCTTC | ATTACGCGCT | TCGGTCGTAC | TATTTTCCAT | CACTTCTTAC | ACTCCCTAGT | 6720 |
| 15 | TTTTTGAAAA | TCAAATGTAT | CATTCGGCTC | GATACAACTT | AACAGTGGTA | AGTCTTCTTT | 6780 |
| | AATAATTTGT | GCAaCAACAT | TCACATTGTC | ATGTGCACTA | AGCGTTTGTC | TCACAATTTG | 6840 |
| 50 | CATTTCGCCT | TGATAACGTC | CGTTATTCAA | ATTATCAACG | GTTACTGAAC | CAATGCGTCG | 6900 |
| 50 | TTGCGTCGTA | AACTGTGGTT | GAATCGAATG | TGGACATATT | TGTCTTGACG | TTTCCGAACG | 6960 |
| | 3 ATC 3 C 3 TOTAL | TCCCC3 TTS T | GGGGGGGT C | ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ | | | |

| | AAGTTGTTTT | GCCTGCCTCA | TTTCAATCAA | TGAGTCTCCA | ACTAACACTT | CAGATACACC | 7140 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | AGTTTCTTGT | AATAATTTAG | CTGCAACGAC | AGGATGACTA | TGTCTCGTTG | CTTCAATTGT | 7200 |
| 5 | TGGCAAGCCT | TTATGCAAAG | GACCTCGCAA | ATCACTCCCT | ACAATAAAAC | CATATATTTG | 7260 |
| | TGCCTTTGGA | TTAAATTGAT | AAATGAGTTC | ATTTTTCTTA | TTGACCAAGT | CAACAGATAA | 7320 |
| 10 | TCCCGTATCT | GGTCTTGGAT | AATAGTTATG | ACAAAATGAA | AGTAATGTAA | AATCATTCAA | 7380 |
| , 0 | TTGTTGATGT | AAGCTTGTTA | ACAATTCCCG | GGAAATAATA | CTTGCATTCA | AACAGCACTT | 7440 |
| | TAAACCCTGT | GCCATTATCG | CTTCGATTGC | CTCAATTGAT | GTACTATGAT | CGATACGAAT | 7500 |
| 15 | CATAAATTGT | GCATCATATT | GTCGAAGATG | GTCATAAAAA | GATGGTGTTA | AAATAGATGG | 7560 |
| | ATTAGCATCT | ATGAGGTAAG | TCACTTGTTC | ATGTTTTAAT | AAATTGAGTA | GTTTTGTGAA | 7620 |
| | ATAATGATAT | TTTGTCTCGT | CATCTTCTTC | TGGTATTTGT | ACAGATGTAA | AAATCATTTG | 7680 |
| 20 | GTAACCTTGT | TTAATCATTC | GCTTAATATA | CGCTTCATCT | AAAGGTTGTC | CTAAATACAC | 7740 |
| | TGAAAAGCCT | GTCAAAGTAG | CCCTCCTTAA | CAATATAATT | ATTAGGAAAA | TATAGTTGAT | 7800 |
| | TTGTGTAATC | GCTTACATTT | TACTATAAGA | GAAAACACAT | TACAATATTA | ATCAGTTAAA | 7860 |
| ? 5 | GCCTGTTCAT | TGTAATAATC | TTACATATTT | CTGTCACAAG | TTAATTATTA | CACCATCAAA | 7920 |
| | GATTATCCTT | TCTTTTAAGT | GCTGATAATA | GCTGCTACTG | CTGGATTATT | ACAATAACTT | 7980 |
| 3 <i>0</i> | TTATACATTT | TATTCAGGAT | TATCTTATAT | TATGTTTTAA | TAATAATCTG | TGAACAATTA | 8040 |
| | AGAGATTTGA | AATTGAATTT | AATAATTGTA | TTGAAAACGC | ATACTTCACC | ATGCTAAAAT | 8100 |
| | AGGAGTCGCA | AACAAATAAG | ATTCAATAAG | ATGTGATGGT | TACCAACACA | GTCTATTTGC | 8160 |
| 35 | TCGTGTCTTT | TTTTATTGAA | TCTTAAATAA | TAAATACAAC | TTTGGAGGTT | GGACAAGTGA | 8220 |
| | GGAAGAAACT | TTTCGGTCAA | TTGCAACGTA | TTGGTAAAGC | GCTAATGTTA | CCTGTTGCGA | 8280 |
| | TTTTACCAGC | AGCTGGTCTG | TTATTAGCTA | TCGGTACAGC | TATGCAAGGT | GAATCATTAC | 8340 |
| 10 | AACACTACTT | GCCGTTTATA | CAAAATGGTG | GCGTACAAAC | TGTCGCTAAA | TTAATGACAG | 8400 |
| | GTGCTGGTGG | TATCATTTTT | GATAACTTGC | CTATGATTTT | CGCATTAGGT | GTCGCAATCG | 8460 |
| | GATTAGCTGG | CGGTGATGGC | GTAGCAGCTA | TCGCAGCATT | CGTCGGTTAC | ATAATCATGA | 8520 |
| 15 | ACAAAACAAT | GGGCGACTTT | TTACAAGTTA | CACCTAAGAA | TATTGGTGAT | CCAGCGAGTG | 8580 |
| | GTTACGCTAG | CATTTTAGGT | ATCCCAACAT | TACAAACAGG | TGTGTTCGGC | GGTATTATAA | 8640 |
| 50 | TCGGGGCCCT | GGCAGCTTGG | TGTTATAACA | AGTTCTATAA | CATTAACTTA | CCATCTTATT | 8700 |
| | TAGGTTTCTT | CGCTGGTAAG | CGTTTCGTAC | CTATTATGAT | GGCTACAACA | TCATTTATTT | 8760 |
| | TAGCATTCCC | AATGGCATTA | ATTTGGCCAA | CGATTCAATC | AGGATTAAAT | GCATTCAGTA | 8820 |

| | TATTAATTCC | ATTCGGTCTA | CATCACATTT | TCCACGCACC | GTTCTGGTTC | GAGTTTGGTT | 8940 |
|------------|------------|------------|------------|------------|------------|------------|-------|
| | CATGGAAAAA | TGCAGCTGGT | GAAATTATTC | ACGGTGACCA | ACGTATCTTT | ATCGAACAAA | 9000 |
| 5 | TTCGTGAAGG | CGCACATTTG | ACAGCTGGTA | AATTCATGCA | AGGTGAATTC | CCTGTTATGA | 9060 |
| | TGTTCGGTTT | ACCTGCAGCA | GCTTTAGCAA | TTTATCACAC | AGCTAAACCT | GAAAATAAGA | 9120 |
| 10 | AAGTAGTAGC | AGGTTTAATG | GGTTCTGCTG | CTTTAACATC | ATTCTTAACT | GGTATTACAG | 9180 |
| | AACCATTAGA | ATTCTCATTC | TTATTTGTAG | CACCATTATT | ATTCTTTATT | CACGCaGTAC | 9240 |
| | TTGATGGTTT | ATCATTCTTA | ACATTGTACT | TATTAGATCT | TCATCTAGGT | TATACATTCT | 9300 |
| 15 | CAGGTGGTTT | CATCGACTAC | TTCTTACTCG | GTATACTACC | TAATAAGACA | CAATGGTGGT | 9360 |
| | TAGTCATTCC | TGTAGGTCTT | GTATACGCAG | TTATTTACTA | CTTCGTATTC | CGATTCTTAA | 9420 |
| | TTGTAAAATT | AAAATACAAA | ACACCAGGTC | GTGAAGATAA | ACAATCACAA | GCGGCTACTG | 9480 |
| 20 | CTTCAGCAAC | TGAATTACCA | TATGCAGTAT | TAGAAGCTAT | GGGTGGCAAA | GCAAACATTA | 9540 |
| | AACATTTAGA | CGCTTGTATC | ACACGTCTAC | GTGTTGAAGT | TAACGACAAA | TCTAAAGTTG | 9600 |
| | ATGTTCCTGG | TTTGAAAGAT | TTAGGCGCAT | CTGGTGTATT | AGAAGTCGGC | AATAATATGC | 9660 |
| ?5 | AAGCAATTTT | TGGTCCTAAA | TCTGACCAAA | TCAAACATGA | AATGCAACAG | ATTATGAATG | 9720 |
| | GTCAAGTAGT | AGAAAATCCT | ACTACTATGG | AAGACGATAA | AGACGAAACT | GTTGTTGTTG | 9780 |
| 30 | CAGAAGATAA | ATCTGCAACA | AGCGAATTGA | GCCATATCGT | GCATGCACCA | TTAACTGGTG | 9840 |
| | AAGTAACACC | ATTATCAGAA | GTGCCTGATC | AAGTGTTCAG | CGAAAAAATG | ATGGGTGACG | 9900 |
| | GTATCGCTAT | CAAACCTTCA | CAAGGTGAAG | TTCGTGCACC | ATTCAACGGT | AAAGTACAAA | 9960 |
| 35 | TGATTTTCCC | AACAAAACAT | GCAATTGGTC | TTGTATCAGA | TAGTGGTTTA | GAACTATTAA | 10020 |
| | TCCACATCGG | TTTAGACACT | GTTAAATTAA | ACGGAGAAGG | CTTTACTTTA | CATGTTGAGG | 10080 |
| | AAGGTCAAGA | AGTTAAACAA | GGTGATTTAT | TAATCAACTT | TGATTTAGAC | TACATCCGCA | 10140 |
| 10 | ATCATGCAAA | GAGTGATATT | ACGCCTATTA | TCGTGACACA | AGGAAACATT | ACAAACCTTG | 10200 |
| | ATTTTAAACA | AGGTGAACAT | GGCAACATTT | CATTTGGCGA | TCAATTATTT | GAAGCTAAAT | 10260 |
| | AATGCTTACT | ATAAACAGGT | GCGTATACCT | TCATAAGGTG | ACGCGCCTGT | TTTTTCTTTG | 10320 |
| 45 | CTATTGTATT | TTGCAGCATC | ATTGATAGTT | CGCTCTCCCC | TTAAATTTTG | AATTTTAAGA | 10380 |
| | TCATCAATTA | AAGCCCCCCT | TCATACTCAT | TTCCTAAAAA | ATATTAATTG | TTCACTATTG | 10440 |
| 5 <i>0</i> | TTAGCGTTTT | CACAACAAAG | TCAACTTCCT | TGACCTTACA | CTATATTCGA | GGCTATCATT | 10500 |
| | TTAAGTGTAA | ATATAGAGAA | AAGGTGGCTT | TTTTTATGAA | ACAACGCATT | GGAGCTTACT | 10560 |
| | TAATTGACGC | TATTCATCGA | GCAGGCGTCG | ATAAAATTT | TGGTGTTCCT | GGTGATTTTA | 10620 |

| | ATGAATTAAA | CGCAAGTTAC | GCAGCGGACG | GTTATGCCCG | TCTTAATGGA | CTCGCTGCAT | 10740 |
|----------|------------------|------------|------------|------------|--------------|------------|-------|
| | TAGTTACTAC | ATTTGGTGTT | GGCGAATTAA | GTGCCGTCAA | CGGTATCGCA | GGTTCATATG | 10800 |
| 5 | CTGAACGCAT | ACCTGTCATT | GCGATTACAG | GTGCGCCGAC | ACGTGCTGTT | GAACAAGGCG | 10860 |
| | GTAAATATGT | ACATCACTCA | CTTGGTGAAG | GTACATTTGA | CGACTATCGA | AAAATGTTTG | 10920 |
| | CACATATAAC | CGTTGCACAA | GGTTATATCA | CACCTGAAAA | TGCAACAACC | GAAATACCAC | 10980 |
| 10 . | GTTTAATTAA | TACAGCAATC | GCCGAAAGAC | GCCCAGTTCA | TTTACATTTA | CCAATCGATG | 11040 |
| | TCGCAATCTC | TGAAATTGAG | ATACCGACAC | CATTTGAAGT | GACGGCAACT | AAATATACGG | 11100 |
| 15 | ATGCATCAAC | ATATATAGAG | TTATTAGCAA | CTAAACTGCA | TCAAGCGAAG | CAGCCTATCA | 11160 |
| | TCATTACTGG | ACATGAAATT | AACAGTTTTC | ACCTCCATCA | AGAATTAGAA | GATTTTGTAA | 11220 |
| | ATCAAACACA | GATACCAGTA | GCACAACTTT | CATTAGGAAA | AGGTGCTTTT | AATGAGGAAA | 11280 |
| 20 | ATCCATATTA | TATGGGTATT | TACGATGGGA | AAATTGCCGA | AGATAAAATA | CGAGATTATG | 11340 |
| | TGGACAACAG | CGATTTAATT | TTAAATATTG | GAGCCAAATT | AACAGATTCA | GCAACAGCAG | 11400 |
| | GTTTTTCATA | CCAATTCAAT | ATCGATGATG | TCGTTATGTT | AAATCATCAC | AATATCAAAA | 11460 |
| 25 | TTGACGATGT | TACAAATGAT | GAAATATCTC | TACCATCATT | GTTAAAACAG | TTATCCAATA | 11520 |
| | TTTCATATAC | GAATAACGCA | ACGTTCCCTG | CGTATCATCG | TCCAACATCA | CCCGATTATA | 11580 |
| | CTGTTGGCAC | AGAACCATTA | ACACAACAAA | CTTATTTTAA | AATGATGCAA | AATTTCTTAA | 11640 |
| 30 | AACCAAATGA | TGTCATCATT | GCTGATCAAG | GTACATCATT | CTTTGGTGCT | TATGATTTAG | 11700 |
| | CATTATACAA | AAACAATACT | TTTATAGGGC | AACCGTTATG | GGGTTCTATC | GGCTATACAT | 11760 |
| 35 | TACCTGCAAC | ATTAGGTTCA | CAATTAGCAG | ACAAAGATCG | TCGTAACTTA | TTATTAATTG | 11820 |
| | GTGATGGCTC | ATTGCAACTA | ACTGTTCAAG | CTATTTCAAC | TATGATTAGA | CAGCATATTA | 11880 |
| | AACCÉGTATT | ATTTGTGATT | AATAATGACG | GCTATACGGT | AGAACGACTT | ATTCACGGCA | 11940 |
| 10 | TGTATGAACC | TTATAATGAA | ATTCACATGT | GGGATTATAA | AGCTTTACCA | GCTGTATTTG | 12000 |
| | GTGGTAAAAA | TGTTGAAATT | CATGACGTTG | AATCATCAAA | AGATTTACAA | GACACGTTTA | 12060 |
| | ATGCAATTAA | TGGTCATCCC | GATGTGATGC | ATTTTGTCGA | AGTCAAAATG | GCTGTCGAAG | 12120 |
| 45 | aCGCACCGAA | GAAACTCATC | GATALCGCTA | AAGCTTTTTC | ACAACAAAAT | AAATAATTTC | 12180 |
| | ATCGTATACA | GGGTATAAGT | TTAAGCGAAT | ACTTTATTAA | ACGAATAGGA | CTCTGATATA | 12240 |
| | AGATGATTAA | TTTTAATAAA | ACCGCTTTAG | TGTTAATCGA | CCTGCAAGAA | GGTATTCTTA | 12300 |
| 50 | AAATGGATTA | TGCCCCATAT | ACAGCTGAAA | ATGTCGTTCA | AAACGCTAAT | AAATTAATAG | 12360 |
| | A WC THEFT THE A | **** | CCCTTTTN | | CCTC N NOTTO | mamcamcoma | 12424 |

| | AGTCGTTTCC | ATCATTTATT | AGACAAGAGA | GATGACGATT | TTGTCATAGA | CAAACGACAT | 12540 |
|----|----------------|-----------------|--------------------|----------------|------------|-------------|-------|
| | TTTAGTGCAT | TTGTAGGAAC | AGATTTGGAC | TTACAATTGC | GACGTCGAGG | AATTGATACG | 12600 |
| 5 | ATTGTTCTTG | GTGGTGTCGC | AACGCATATT | GGCGTAGATA | CGACAGCGCG | AGATGCCTAT | 12660 |
| | CAATTAAACT | ACAATCAGTT | TTTTGTTACA | GATATGATGA | GTGCACAAAA | CGAAACGCTA | 12720 |
| | CATCAATTTC | CAATAGATAA | TGTATTCCCA | TTGATGGGAC | AAACAATAAC | TACAAACGAC | 12780 |
| 0 | TTTCTAAATA | TATTGAACTA | AACATATACT | TCCCCCCTTC | GATCATGTTG | AGGGGGATCT | 12840 |
| | TTATTTCACA | AAGTATTAAT | ACGTCGGGTT | GTCTAACCTT | CTATATTTAA | CATATTCTAT | 12900 |
| 5 | ATCTGTTAAA | TCGTTCTTAA | CTTACGCCCC | TACTACATAA | AAAACAGTAT | TTATTCCGGA | 12960 |
| 3 | ATTTTCAAAA | AATTTAGTAT | TTATTGCAAA | ATTATGTATC | ACTTTATGTT | TAATTTTTGA | 13020 |
| | TATTATCTTA | ATTAAGTAGA | TTTTTATAAG | TTCTAAAAAG | GAGAACAAAT | ACATATATGA | 13080 |
| 0 | AGAAGAAACT | AACATTTAAA | GAAAACATGT | TTATAGGTTC | TATGTTATTT | GGTTTATTCT | 13140 |
| | TTGGTGCCGG | CAATCTTATC | TTCCCAATAC | ACTTGGGTCA | AGCTGCTGGT | TCTAACGTTT | 13200 |
| | TTATCGCTAA | CTTAGGATTT | TTAATTACAG | CAATTGGCTT | ACCATTTCTA | GGTATCATTG | 13260 |
| 25 | CTATTGGCAT | TTCAAAGACA | TCTGGTTTAT | TTGAAATTGC | ATCGCGTGTT | AATAAAACAT | 13320 |
| | ATGCTTACAT | TTTCACGATT | GCCTTATATC | TAGTTATCGG | ACCATTTTTC | GCCTTACCTA | 13380 |
| | GACTGGCAAC | GACATCATTT | GAAATTGCAT | TTTCGCCATT | TTTATCACCA | AAGCAAATCA | 13440 |
| 80 | CTTTATATTT | ATTTATTTTT | AGCTTCGTCT | TCTTTGTGAT | TGCATGGTTT | TTTGCGAGAA | 13500 |
| | AGCCATCAAG | AATTTTAGAA | TATATCGGTA | AATTTTTAAA | TCCGGTATTC | TTAGTATTAT | 13560 |
| | TAGCAATTAT | TTTATTATTT | GCTTTTATCC | ATCCATTAGG | TGGCATATCT | GATGCACCTA | 13620 |
| 35 | TTAGTAAACA | ATATCAATCA | CATGCCTTAT | TTAACGGCTT | TTTAGATGGA | TACAATACCT | 13680 |
| | TAGATGCGCT | AGCGTCATTG | GCATTTGGTA | TTATCATTGT | TGCAACGATT | AAAAAGTTAG | 13740 |
| 10 | GTATCGAAAA | TCCAACTGAT | ATCGCTAAAG | AAACAATTAA | GTCTGGTACT | ATCAGTATCA | 1380 |
| | TTATGATGGG | GATCATTTAT | ACCCTACTAG | CAATCATGGG | TACATTAAGT | ATTGGTCATT | 1386 |
| | TCAAACTTAG | TGAAAATGGT | GGTATTGCCT | TAGCGCAAAT | TACTCAATAC | TACTTAGGTA | 1392 |
| 15 | ACTACGGTAT | CGTCCTGTTG | TCACTTATCG | TTATGGTTGC | TTGTTTAAAA | ACAGCCATCG | 1398 |
| | GTTTGATTAC | GGCATTTTCA | GAAACATTCG | AACACCTTTT | CCCTAAAATG | AATTACCTAG | 1404 |
| | CGATTGCAAC | AGTTGTAAGC | TTTATTTCGT | TCTTATTCGC | GAATGTTGGT | TTAACTAAGA | 1410 |
| 50 | TTATTATGTA | CTCAGTCCCA | GTGTTAATGT | TCTTATATCC | ATTAGCAATT | GCCTTGATTG | 1416 |
| | T2 CT2 2 C2 TT | 3 mmn 3 cm 3 cc | 3 3 3 TOTO CO 3 TO | DOTTON DE DOTT | TATTATCA A | TOTA COATTO | 1422 |

| | GCACATCATT | CTCACAAACT | TTGATTAATT | TCAGCCAAAA | ATATTTACCA | TTATCAGACA | 14340 |
|-----|------------|-------------|------------|------------|------------|------------|-------|
| | TTGGTATGGG | CTGGGTTGTT | CTCAGTTTGA | TTGGTTTCAT | TATCGGCTTC | ATTATTTATA | 14400 |
| 5 | AAATTAAGCA | TCGTAAAATT | CCACAAGCAT | AATACTATGC | CACAGTCATA | TGTTAAACAT | 14460 |
| | ATGCTTGTGG | CATTTTTTAT | TCATACTACA | TTAAACTGCA | ATCGTATACA | TACATATCAA | 14520 |
| 10 | TGATTATCCA | CAAAAAAATAT | TAGTACTTTC | ATTTTACAAA | TCACATTAAT | ACAAACACAA | 14580 |
| ,,, | CCTTATCTTT | ATATTATTAA | ATTTATATTT | GACACTTATA | TTGAACAACT | GTAATATATT | 14640 |
| | AATATTAATT | CTTTAAAATG | TATAAATATA | AAGGAGGGAG | ACCGATGAAT | TCAATCATTG | 14700 |
| 15 | aATTAACTGA | TTATTATAGC | TCTAATAATT | ATGCACCACT | TAAGCTTGTC | ATTTCTAAAG | 14760 |
| | GTAAAGGTGT | CAAAGTTTGG | GATACTGATG | GCAAACAATA | TATAGATTGC | ATTTCGGGTT | 14820 |
| | TTTCAGTTGC | AAACCAAGGC | CATTGTCATC | CAACAATTGT | TAAAGCGATG | ACAGAACAAG | 14880 |
| 20 | CTTCAAAGTT | GTCTATCATT | TCACGTGTCC | TTTATAGTGA | CAATCTCGGG | AAATGGGAAG | 14940 |
| | AAAAAATTTG | TCATCTTGCT | AAGAAAGACA | AAGTACTCCC | CCTTAACTCT | GGTACTGAAG | 15000 |
| ÷ | CTGTTGAAGC | AGCCATTAAA | ATTGCTAGAA | AATGGGGCTC | TGAAGTTAAA | GGCATTACTG | 15060 |
| 25 | ACGGACAAGT | TGAAATCATC | GCTATGAATA | ACAATTTTCA | CGGTCGTACA | CTTGGCTCAT | 15120 |
| | TATCACTATC | TAACCACGAC | GCATATAAAG | CAGGATTTCA | CCCCCTACTT | CAAGGCACTA | 15180 |
| | CAACAGTAGA | TTTTGGAGAC | ATTGAACAAT | TAACACAAGC | TATTTCACCG | AATACAGCAG | 15240 |
| 30 | CAATTATTTT | GGAACCAATT | CAAGGTGAAG | GTGGCGTTAA | TATACCACCG | AAAGGATATA | 15300 |
| | TTCAAGCTGT | GCGTCAACTA | TGTGATAAAC | ATCAAATATT | ATTGATTGCA | GATGAAATTC | 15360 |
| 35 | AAGTTGGTCT | TGGTAGAACT | GGGAAATGGT | TTGCTATGGa | ATGGGAGCAA | GTCGTTCCAG | 15420 |
| | ACATTTATAT | TTTAGGTAAG | GCATTGGGTG | GCGGCTTATA | CCCTGTATCT | GCTGTACTTG | 15480 |
| | CAAATAATGA | TGTCATGCGT | GTTCTAACAC | CAGGTACACA | TGGTTCAACA | TTTGGTGGTA | 15540 |
| 40 | ACCCTTTAGC | CATTGCAATA | TCGACGGCAG | CGCTTGATGT | ACTTAAAGAT | GAACAACTGG | 15600 |
| | TTGAACGATC | AGAACGCTTA | GGTTCATTTT | TATTAAAAGC | GTTGCTACAA | CTTAAACATC | 15660 |
| | CTAGTATTAA | AGAAATTAGA | GGTCGTGGTT | TATTTATAGG | CATAGAGCTT | AACACAGATG | 15720 |
| 45 | CTGCACCTTT | TGTGGATCAA | CTGATTCAAC | GTGGAATCTT | ATGCAAAGAC | ACGCATCGTA | 15780 |
| | CTATCATTCG | ATTGTCTCCA | CCTCTAGTCA | TTGATAAAGA | GGAAATCCAT | CAAATTGTTG | 15840 |
| | CAGCTTTTCA | AGACGTTTTT | AAAAATTAAC | AATTAATCAT | TTATATATGA | CATAGGAGGG | 15900 |
| 50 | ATTCATGATG | attaaagtag | GTATCGTTGG | CGGTAgcGGT | TATGGCGCAA | TTGAATTAAT | 15960 |
| | TCGATTGTTA | CAAACACATC | CTCATGTAAC | GATTGCACAC | ATCTACTCAC | ATTCAAAAGT | 16020 |

ACTTACAGTG GATAATAATG ACTGTGATGT AATTTTCTTT GCGACACCAG CACCCGTAAG 16140 TAAAACATGT ATCCCTCCCT TAGTAGAAAA AGGTATTCAT GTTATCGATT TATCTGGCGC 16200 ATTTAGAATT AAGAATCGTG AAATATATGA AGCATATTAC AAAGAAACTG CTGCAGCACA 16260 AGATGATTTG AATCATGCTA TTTACAGCAT TTCAGAATGG CAATCGTTTG ATAACAATGG 16320 AACGAAGCTC ATTTCTAATC CTGGCTGTTT CCCTACAGCA ACATTATTAG CATTACATCC 16380 10 ACTTATTAGC GAAAAAATAG TAGATTTGTC ATCTATTATT ATTGATGCTA AGACCGGCGT 16440 GTCAGGTGCT GGTCGTTCAT TATCACAACG GTTCATTTTT CAGAAATGAA TGAAAATCTA 16500 AGCGCTTATG CAATCGGAAA CCATAAACAC AAACCGGAAA TCGAGCAATA TTTATCTATC 16560 15 ATTGCGGGTC AAGATGTATC AGTCATATTT ACACCACATC TCGTACCAAT GACACGAGGT 16620 ATTITATCAA CAATATATGI CAAATTATCA TCTGAATATA CGACTGAATC ATTACATAAA 16680 TTAATGACCT CTTATTATGC TAATCAGCCA TTTGTCAGAA TTAGAGATAT TGGGACTTTT 16740 20 CCAACCACAA AAGAAGTACT CGGTAGTAAC TACTGCGATA TCGGCATCTA TGTAGATGAA 16800 ACAACGCAAA CAGCAATTTT AGTATCAGTG ATTGATAACC TTGTCAAAGG CGCAAGTGGG 16860 25 CAAGCCATTC AAAATTTAAA TATATTATAT GATTTTGAAG TGACGACTGG CCTAAATCAA 16920 TCACCAGTTT ATCCATAAGG GGTGTTAGAA TGAAACATCA AGAAACGACA TCACAACAAT 16980 ATAACTTTTC AATTATTAAA CATGGCGATA TCAGTACACC TCAAGGCTTC ACGGCTGGTG 17040 30 GTATGCACAT CGGTTTACGC GCTAACAAAA AAGACTTTGG GTGGATTTAC TCATCGTCTT 17100 TGGCAAGTGC AGCTGCCGTA TATACTTTAA ATCAGTTTAA AGCTGCACCA CTTATTGTCA 17160 CTGAAGACAC TITACAAAAG TCTAAAGGAA AATTACAAGC ACTTGTTGTT AATTCAGCTA 17220 35 ATGCAAATTC TTGTACCGGT CAACAAGGCA TAGATGATGC ACGACAAACA CAAACATGGG 17280 TTGCTCAACA ACTTCAAATA CCATCTGAGC ATGTTGCTGT TGCTTCAACT GGGGTCATTG 17340 GTGAATATTT GCCTATGGAT AAAATTAAGA CTGGGACCGA ACATATTAAG GATGCTAATT 17400 40 TTGCAACGCC AGGTGCGTTT AACGAGGCAA TTTTAACAAC TGATACCTGT ACAAAACATA 17460 TCGCTGTATC ACTAAAAATC GATGGTAAAA CCGTTACAAT TGGTGGTAGC ACCAAAGGTT 17520 CAGGTATGAT TCACCCAAAT ATGGCTACCA TGCTTGCTTT TATAACAACC GATGCATCGA 17580 45 TTGAATCGAA TACACTTCAT CAATTATTAA AATCTTCGAC TGACCATACA TTTAATATGA 17640 TTACTGTTGA TGGCGATACA AGTACAAATG ACATGGTATT AGTCATGGCA AATCACCAAG 17700 50 TTGAACACCA AATACTTAGT CAAGACCATC CACAATGGGA AACATTTGTT GATGCATTCA 17760 17820 ATTTTGTCTG TACATTTTTA GCTAAAGCTA TAGCCAGAGA TGGCGAAGGC GCAACAAAGT

| | CTATCGTAAG | TTCAAATCTA | GTAAAATCAG | CTATTTTTGG | CGAAGATGCC | AATTTTGGTC | 1794 |
|-----|------------|------------|------------|------------|------------|------------|-------|
| | GAATCATTAC | AGCTATTGGC | TACAGCGGAT | GTGAAATTGA | TCCTAACTGC | ACATATGTTC | 18000 |
| 5 | AACTGAACCA | AATACCTGTC | GTTGATAAAG | GTATGGCTGT | ACTATTTGAT | GAGCAAGCTA | 18060 |
| | TGTCGAATAC | ATTAACTCAT | GAAAATGTCA | CAATTGACGT | TCAGCTTGGT | TTAGGTAACG | 18120 |
| 10 | CTGCAGCGAC | TGCATACGGT | TGTGATTTAT | CCTATGATTA | TGTGCGTATC | AACGCATCAT | 18180 |
| 10 | ATCGAACATA | AGGTGGTGTT | GGTTAGATGA | AATTTATTGT | CATTAAAATT | GGTGGCAGTA | 18240 |
| | CACTTAGTGA | CATGCATCCA | TCAATTATTA | ACAACATTAA | GCATTTACGA | TCAAACAACA | 18300 |
| 15 | TCTACCCCAT | TATCGTTCAT | GGCGGTGGCC | CATTTATTAA | TGAAGCATTA | TCAAACCAGC | 18360 |
| | AAATCGAGCC | ACACTTTGTT | AATGGCCTAA | GAGTGACTGA | TAAAGCAACC | ATGACCATTA | 18420 |
| | CTAAACACAC | GCTCATTGCA | GACGTTAACA | CTGCATTAGT | AGCTCAATTT | AACCAGCACC | 18480 |
| 20 | AATGTTCTGC | AATAGGCTTA | TGTGGTTTGG | ATGCACAGCT | GTTTGAAATT | ACATCTTTTG | 18540 |
| | ATCAACAATA | TGGATATGTC | GGTGTTCCGA | CCGCTTTAAA | TAAGGATGCT | TTACAGTATT | 18600 |
| | TATGTACTAA | ATTTGTACCT | ATCATCAATT | CGATTGGTTT | CAATAACCAT | GATGGAGAAT | 18660 |
| 25 | TTTACAATAT | TAATGCTGAC | ACGCTTGCCT | ATTTTATTGC | ATCATCATTA | AAAGCGCCTA | 18720 |
| | TTTATGTATT | AAGTAATATT | GCAGGTGTAC | TCATCAATGA | TGTTGTTATA | CCTCAATTGC | 18780 |
| | CATTAGTCGA | TATTCATCAA | TATATTGAAC | ATGGTGATAT | TTATGGAGGT | ATGATTCCCA | 18840 |
| 30 | AAGTGCTAGA | TGCCAAAAAT | GCGATTGAAA | ATGGCTGTCC | TAAAGTTATC | ATTGCATCAG | 18900 |
| | GAAACAAGCC | AAATATCATT | GAATCTATTT | ACAATAATGA | TTTTGTTGGC | ACAACAATCC | 18960 |
| 35 | TTAATTCATA | ACTATGAAAT | TAAGGCCTAA | CAAGTTTTGA | CACGCGAGAT | GATTCCAGTT | 19020 |
| ,,, | CGATTATCCA | TTGCGCTAAA | ACATTTATTT | ACCGTTCATC | TCGTTAACAA | TTTTGAATAC | 19080 |
| | AGTACGATAC | AATATGAGAT | GTAAAAAACT | AATAACCTTT | TACAAATTTG | TTTATCAAAA | 19140 |
| 10 | TATTTTAAGT | TTTGCAAAGC | TTTTTATTGT | GATTATTTTC | ACAAAATACT | ATAATGAGGA | 19200 |
| | TAGTAAATAG | AGAGGAGTCC | TTAAGTTGAC | GAAACGACAA | ATGGGTATAT | TCATTTATGC | 19260 |
| | TGGAATTATC | GGTGGCTTGT | TATCTGGAAT | TGTAAAATTA | GGTTGGGAGG | TCATGTTTCC | 19320 |
| 15 | ACCTCGCACA | CCAGAACGTA | ATGCAACGAA | CCCACCTCAA | GAGTTATTGC | AACAATTAGG | 19380 |
| | ATTTAGTAGT | GAGTTTACGC | ATCAAACATA | TACATTTTCA | AATATGGAAT | TGCCTTGGGT | 19440 |
| | AAGCTTTATT | GTCCACTTTA | GTTTTTCTAT | CGTCATTGCA | ATTATTTACT | GCATATTAGT | 19500 |
| 50 | TAAAAAATAC | GCTTACTTAG | CAATGGGACA | AGGTGCTGTT | TTTGGTATTG | CTATTTGGGT | 19560 |
| | ATTATTCCAC | СТТАТСАТТА | ТСССДДТСДТ | GCATACTGTA | ССТСТСТСТ | CCCATCAACC | 19620 |

| | AGTGCGACAA | CATTTTGTCT | ATCGCTATAA | ATTAAATTAA | TACACTGACT | AACATTAACG | 19740 |
|------------|------------|------------|------------|------------|------------|------------|-------|
| | TGAGTTTLAA | ATCATCGTTT | GAGTATGATG | ATTGATGCTC | ACGTTATTTT | ATTAACTGAC | 19800 |
| 5 | ATGATATGAT | TCCAGCCAAC | TTACGTGAGC | ATTAAAGTCT | CAAATGCGTC | GTAACAAACT | 19860 |
| | ATTATTTTCG | GTAATTTCAA | TATTGCTCAG | TATATTTTTA | CCTTATCACT | TACTTTAATC | 19920 |
| | TCGTCATGAT | TTTGAATGAT | GCCATCGTGT | ATTCACCTTT | CATTTTTCCA | ATAAAAAAAC | 19980 |
| 10 | ATCTAACAGT | AAACATTTAG | GCAGTATAGT | TTAAAATCAC | TGCGCAATGA | TACTGTCAGA | 20040 |
| | CGTCATATTA | ACTACTCAAT | AACTGAAATA | CAGACACTTT | TTTATAACCC | CAGGGTGCCT | 20100 |
| 15 | GTCCTAAGAA | ACATACCTGT | ACCATAAACT | GATCAAAAAT | AAATTGTTTG | AACTTCACTT | 20160 |
| | CACGTGATTG | ATAAAAGTGT | GATTGTGTCA | TATCATAAAT | GTCCAATCCT | TTGATTAAAC | 20220 |
| | CTTCACCAAT | CAATTTTGTA | AAACTTTCTT | TTTGTGTCCA | TATTTGATAA | AAATCATTTA | 20280 |
| 20 | AACTACATAT | TTGATGTGCT | TCGTTTGTAG | AGAAACACGT | CACTAACGTA | CGCCAGTCTA | 20340 |
| | AACGTTGTGA | TATCTTTTCG | ATATCAATAC | CAACTGGTTC | TTTATCGACA | ACACACACGA | 20400 |
| | TATAAGGATA | ACTATATGAT | AAGCTCACAT | AGATGGGCTG | TCCATCACGA | TTGTGTTGAA | 20460 |
| 25 | CAATATCTGC | CTTACCTCGT | GGCGAAATGT | GATAATGCCA | TTCATGTGGT | AATAAACCTG | 20520 |
| | TGTCATGTTG | AATTCCATAT | TGCACTAAAA | TATCTCCCAA | TCTGTGCATG | AGTTTATCTT | 20580 |
| | GATTGTATCT | ATAGTTGACT | GTACGCGGTT | TTTTATATGA | CCAACGACTT | TGTGATATTA | 20640 |
| 30 | ATTCTTCAAT | ACTTTTCAAG | TTACTCTGTA | ATTGCATTAC | AAATACTGTC | ATAACTTTCC | 20700 |
| | CTACTTACTT | ATTGAATATT | GTTTTGATAT | ATTGTGCCCA | ATGATACAGC | CAATTGTTAG | 20760 |
| 35 | TTATCGTTGG | CCATTTTTCA | CTGATGTGAT | TCATTATTTT | TAATGTTAAT | GTTGTATCTA | 20820 |
| J J | TCATTGCTAG | TIGTTGTTCA | CGGTCAACAC | TAGTTAATCC | AATCGTTTCG | TACATGTCTT | 20880 |
| | GTTTCTGTAA | AATTTCATTA | AATGATTCAT | CGCTGACGAG | TTCAATTTCT | TTGCGCTTAA | 20940 |
| 40 | CGCATTCTAA | CAAAGATTTC | ACCGGCATIT | TATTAGGTGA | TAGCACATGG | TAAATGATTT | 21000 |
| | GTGGTGTGTT | GACCTGTGCT | AATGCGACAA | TTTGTCTTGC | AGTCGTATCC | ACAAAAGAAA | 21060 |
| | AATCTACAGG | CATTTCAGCC | ATGCTAACCC | CGATACAATC | CAGTTGTAAC | AAATCATTCA | 21120 |
| 45 | TTACCATTGA | AAAACGGTTA | GTCTTTATAT | TTCTCATATG | CCATCTTCCA | TTGTAAGGAT | 21180 |
| | TCGTCAAATT | ACCAACACGT | ACAATCCGAC | CATCTAAGCC | ATTATTTACA | GCTTCTAATA | 21240 |
| | CTTTTAATTC | ACTATAAAAT | TTGCTCCGTG | TATATGGTGA | TGTTAGTAGT | TGCCCTTTAT | 21300 |
| 50 | AGACATCCGC | TTSTGAAAAT | GTCACATCTT | CTGTGTCTAT | ATCAAAATAA | GTTCCCACAC | 21360 |
| | TTATCGTAGA | CACATATATT | AACCTTGCAT | GATGTTGTTG | TGCCAAACGT | ATGACATCAA | 21420 |

| | CACCTGCATG AATAATCGTA TCCATGTTTT CTGGTAAAAC AACATCATCC ATACACTCGA | 21540 |
|----|---|-------|
| - | AATCACCAAC AATGACTTCA ATGTTTGATA ACATTATTTC AACCGTCTCT TCTGAAAAAT | 21600 |
| 5 | AATCATTTAA ATTCGTCATC AACTTATACC ATGCTATTTC CTCATTATCA GCACGTATGA | 21660 |
| | AACAATAAAT GCGATGACTG TATCCTTGTA GTACTTCAAT CAGATAAGCA CCTAAAAAAC | 21720 |
| 10 | CTGTCGCGCC AGTCAATAGT GTATTTCCTA GAGGTCGATG ACTTAGACTA TCCTCTAAAA | |
| 10 | TACCCAAGTT ATAACGAGAC ATAACAATCT TTTGTAATTC CGAAAGATTA TCCGGTAATG | 21840 |
| | CAACTAATGA TTGTTGATTT TGGTACATAT AATTAACAAT CTGTCGCACG GTTTTATATT | 21900 |
| 15 | GGTATAATGT CTGCATTGAA ATATGATGGC CAAATCGTTT TAAATGCGAG ACAACTAACA | 21960 |
| | TCGCCTCTAA TGAGTTACCA CCAAGTTCAA AGAAATCATC GTCAACACCG ACATCATTTT | |
| , | GTTTCAATAC CTCTCCAAAT ACATCAACAA ATGTCTGCTC AATTTCATTA GAGGGTTCGC | |
| 20 | TATACACTTT ATTAGACTGT TGTATAGGTG ATGGATTTGG CAAACGCGTA GTATCCACCT | |
| | TGTCATTCGT GGTTAATGGC ATACAATCGA TATGCGTTAT AGTCTTAGGA ATCATATACT | |
| | TAGGCAGCTG ATCATTTAAA TATTGCTTCA AATCCTGTTC CACTTGTTGC TCTCCGACAT | |
| 25 | AATAAGCATT CAATATATCA TGCGTATCAA AGTGACTTAC TGTTACAACA CAATCAGATA | |
| | TACCACGAAT AGCTAATATT GCATTITCAA TITCATCAAG TTCAATACGG TACCCGTTAA | |
| | CTTTCACTTG TTTATCTATT CTTCCTAAAA ATTCAATTTG ACCATCAGAT GTATAACGTG | 22440 |
| 30 | CTAAATCACC ACTATGATAC AACTTTCCTT TACCAAATGG ATTATTTTGC CATTTATCAG | 22500 |
| | CCATTAATTC TGGACGATTA ATATATCCTA TCGCTAAACT ATCACCTGCA ATACACAACT | 22560 |
| 45 | CGCCTGGCAT ACCAATACCG CATAACAAAC CATCTGACAT AATATACACT TGGATGTTAG | 22620 |
| 35 | ATAAGGGTTT GCCAATTGGA ATCGTCTCAG GTATCAAATC ACCACAATGA TGTGACCAAT | 22680 |
| | ACGATGTGAT GACTGTTGAC TCAGATGGTC CATAGGCATT GAAATACGTG CCACAATGCT | 22740 |
| 40 | TCTCAATATA TTTAACAAAG GATGCCGTAC TAGTTGCCCC GCCTGTAATC AACTTTTCAA | |
| | TATAAAAGTC TTCCATAACA CTACACATCT GTAACGGAAT CGACGCAACC GTCACACGAT | 22860 |
| | GCTTATTAAT GAGTTGTTGT AACTGTTCTG GATTAACACG TTCCTCTCTA TCTGGAATCA | |
| 45 | CAAGCGTATG ACCATTTAAC AAACAACAAT AAATCTCCAT AACTGATGCA TCAAAAACAA | |
| | TATTTGCATG TTGCAAAAAT ACTTCATTGT CGCCTAATTG CAATTCAGTT GACCATGCAT | |
| | GCACTAAATT CAACAAATTT CGTTGTCGTA TGGCAACCCC TTTAGGCATC CCGGTCGTAC | |
| 50 | CAGATGTGTA AATAGCATAC ATCTCATTAT CTAACATCGC TGTGTTTTCA AGTTGATTGC | |
| | CATGTAAATC ATCATATTGT TCATTTTCCT TTGATTCAAC AAAGCCTTTA GCATTTTCCA | |

| | TAGCATCCTC | CAAAATTGCA | CCTTGTCGTT | TATTCGGAAA | ATCAATATCG | ATAGGTATAT | 23340 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | AAGATGCACC | TACTTTAACT | GTCGCCAACA | TCGCCGCAAT | CATTTCAAAA | CTACGTTCTG | 23400 |
| 5 | TAAACAAGGC | AACCCGTTGA | CCATTGCCCA | CACCATTTGA | TAGGAGCATG | TGCGCAATGG | 23460 |
| | CATCCACATA | GTTGCGTAAT | GTTTCATACG | TCATTGTCAA | ATCATTCATG | ACTAGCGCAA | 23520 |
| | CATGATTACC | TTGTCGTGAG | ACAACTTCAT | TAAAGTAACT | TATGATAGAT | TTATTTCCCG | 23580 |
| 10 | GGACATTAAG | CATTCGATCG | TTAACATGCG | TATTGACCCA | ATTTAGAAGT | TCCTCCGTGC | 23640 |
| | CGTTTGGTAT | ATCACAAATT | TGTAGTGTAT | CTTGATGCTT | СААААТАТАА | TCAATCATAA | 23700 |
| 15 | TCATACATTG | ATTACCCATG | TGACGAACTG | TTTCTGAGTG | ATATAAATCG | GTATTATACT | 23760 |
| | CGATATTGAT | TGTATAGTCA | TCGCGATCTT | CTTCAATGAT | GAAAGATAAA | TCAAATTTCG | 23820 |
| | CCGTCACTGA | TTTGGGTTGA | ATGTGTGTTA | ATTTACTATG | CCCAAAATGA | GCATGATTCG | 23880 |
| ?0 | TTTCATTGTT | TTGTAGTACT | AACATGACAT | CAAATAATGG | ATTCCGTGAG | GCATCATGTG | 23940 |
| | ATTGATCTAA | GTCATTTACT | AAACATTCGA | ATGGGTATTC | TTGATGCTCG | TATGCCTCCA | 24000 |
| | AACTCATTTC | CTTAACCTCT | TGTAAAAACT | GTGTCCACAT | TTTATCAGGT | GACGGTTGCC | 24060 |
| ?5 | CTCTATATAC | CAACGTATTA | GCAAACATGC | CTAGCATTTG | CTCCGTGCCT | TTATGCATAC | 24120 |
| | GCGCACTCAT | CACACTACCG | ACAACAACAT | CATCTTTTCG | AGCATATCTA | CTTAACAACG | 24180 |
| | TCATGACCAC | ACTCATAAAG | AACATAAAAT | CAGTAATTTG | ATGCTTTTCT | ACATACTTTT | 24240 |
| 80 | GAAGTAGCTG | TCTCATTTGT | TGATTCATTG | TAAATGACAT | CATTGCTCCA | TTTGTCGTTT | 24300 |
| | TAATATTTGG | TCTAACATAG | TCTGTCGGTA | AGCTTAAAAT | AGGTACTTCA | TCTTTGAATT | 24360 |
| | GAGATAACCA | ATATTGTCTA | TGTTTCGTCA | TATCACGATG | CGACATCCAC | TCACTATAGT | 24420 |
| | CTTTATATTG | CAATTTAAGT | GGTAACAATA | ATTTATGTTG | ATAAAGTGCG | TTAAGATCAT | 24480 |
| | TCATTAATTG | TATATTACTC | ATACCGTCAT | TAATGATATG | ATGCGTATCT | ATAAAGAGGT | 24540 |
| 10 | ATGCATGTAA | GGGACTTCTA | ATGTATCTCA | CTCTAATTTG | ACTTGGCTTT | TCCAAATTAA | 24600 |
| | AAGGTGCTAC | AAATTGGCGC | ATGATTTCTT | GTTCATCCGT | AAAATGCGTG | TTAACTTCTT | 24660 |
| | CAAAGTCAAC | TGCAACATCT | GCCACAATAC | GTTGTCGAAC | CTCATCATCT | ACAACAATAT | 24720 |
| 15 | ATTGTGTTCG | TAAAATCTCA | TGTCGCGCTA | TCAAACGCTG | CACTGCTTGT | CGCAATTGAG | 24780 |
| | CTACATTAAG | TTCTGATGAT | AACCGCCATA | AAAAAGGTAC | GTTATACACC | GTATCTTTAT | 24840 |
| | GGTTTGATTT | CCATAATAAA | TACATACGCT | TTTGTGCAGA | GCTCAGCACA | TAATCATCTT | 24900 |
| 50 | TAACTATAGT | TTCTGGAATC | ACTTCATAGT | TTTGTTCTTG | AACCTTAGCA | ATCGCTTGTG | 24960 |
| | CTAGTTCAAA | TACAGTTGGC | TTTTGTAATA | AATCACCAAT | TTGTAATCGT | TTCCCAGTAG | 25020 |

| | AATTATCATG | AATACCTACT | TGATTCACAT | GTAAAATATC | TGCAAAAATT | TGGCATAGCA | 25140 |
|------------|-------------------|------------------|------------|--------------|--|------------|-------|
| | AGTGTTCGGT | ATCTGTACTC | GGTGCTACAT | AGGCATCCGT | ATCGACATAG | TCCATGATAG | 25200 |
| 5 | GCAATGCCTT | CTTATCTAAT | TTCCCATTAA | TAGTAATAGG | AATTTGCTCA | ATATGCATGA | 25260 |
| | AATTAACTGG | TATCATGTAC | TCCGGTAAGG | TCATACGTAA | TTGTGATTTA | ATCTTATTAT | 25320 |
| 10 | GTGATAATGT | ATGCATCGCT | TCATAATAAG | CAACGATATA | CTGATCTTGA | TCATGATTTT | 25380 |
| 10 | GAACAATAAC | AACTGCTTTA | TTAATACCTT | GTATACGCTC | GAGCGCATGC | TCAACCTCTG | 25440 |
| | ACAACTCAAT | CCTAAACCCT | CGAATCTTAA | CTTGTTTGTC | CTTTCGATAT | AAATAATCTA | 25500 |
| 15 | TGTTGCCATC | GGGTAACAAA | CGAACGATAT | CACCACTTCT | ATACATCAGC | TGATTTATAT | 25560 |
| | TTGAATCTTT | GATAAATTTA | TCTGCTGTCA | ATTCTGGCTG | ATTTAAATAA | CCTGCAGCTA | 25620 |
| | ACCCAAAGCC | ACTTGTACAT | AATTCTCCAG | GAATACCAAC | GCCACACCGA | CGCTCGCCTT | 25680 |
| 20 | GCATGATATA | AACATGAGTA | CCCAGAATCG | GTTTACCAAT | AGGAATACGA | TTTGGAACTT | 25740 |
| | TGTTAGGTAT | ATTATACGTC | GTTGTAAATG | TTGTATTTTC | AGTTGGTCCA | TAACCATTAA | 25800 |
| | TAATTTGAGG | ATGCTTCGGT | TTTTGATTAA | GCAAATCCAC | CCACTTAGCA | TTCAATACTT | 25860 |
| 25 | CTCCACCAAT | TAATAAATAC | TTTAACGGTA | CCAATACTTC | TATTCGTTCA | CTAGCAATCT | 25920 |
| | GATTAAATAA | TGAGGAGGTT | AACCACATAG | TATTAACGTC | ATTTTCATTG | ATTAATTGTT | 25980 |
| | CTACCGCTAT | TGGATTTAAT | AATTGTTCTT | TTTTAGCAAC | AATCAGCTTT | CCACCATTGA | 26040 |
| 30 | GCAATGCACC | ATATATTTCA | AATGTTGCAG | CATCAAAGGC | TATAGTTCCT | GATAACAAAA | 26100 |
| | TCGTCGTCTC | TTCATTTAAT | GGTACATAAT | GATTTTGATG | GACCAAGCGA | ACAATACCTC | 26160 |
| 35 | GGTGCGGAAT | TAGTGTCCCT | TTAGGGTTAC | CAGTTGTCCC | CGACGTGTAA | ATAACATAAG | 26220 |
| 33 | CATGATCTTC | TAACGTGTTA | CATTTAGAAA | GATTATCAAT | ATTTTTCCAC | GCTATCTTAT | 26280 |
| | TCAAATCAAT | GTGATTAATA | TTTTGTTTAC | CATTTTCATA | TAAAGCTTGG | TACGTTATTA | 26340 |
| 40 | CAACTTTAGG | CGTTACATCT | TTTAAAATGT | ACTCCTGACG | ATCACTTGGA | TAGTTCGGAT | 26400 |
| - | CAATTGGCAC | GTAAGCCCCA | CCAGCTTTCA | ACACACCTAT | CATCGCTATT | ATCATCTCAA | 26460 |
| | TACTTTTTC | AGCTATGACA | GCGACACGAT | CATTAGGTTC | AACACCATAC | TGGTTTCTCA | 26520 |
| 45 | AACGGTGTGC | TAAATCATTC | GCGCGTGCAT | TCAATGTTTG | ATATGTTATA | AACACTCCGT | 26580 |
| | CAAATTGCAC | AGCGACATGA | TTCGGCGTTG | CTTCAACTTG | TTGCTCAAAT | AAGGTAACAA | 26640 |
| | CTGTTTGCGC | ATCATCTATC | TCAGGCAAAC | TTAAATTGAT | ATCGTCATAT | AATTGAATAT | 26700 |
| 5 <i>Ò</i> | CACGTTCTGT | CATCAAATTA | AGTTCATCTA | CAGTTGTTCG | TTTATTTCCA | TTTTCTTCAG | 26760 |
| | ## * ############ | 1.001.1.001.0000 | CC22002200 | Ch CTCh CCCT | ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ | 1001110010 | 06006 |

| | | | | | | | _ |
|----|------------|------------|------------|------------|------------|------------|-------|
| | CTAAAGATGT | ATGTGCATCA | TGTATTTGAT | GTACATCCTC | AATAACATCA | TTACAACAAG | 26940 |
| _ | ACATCATATG | ATGATAACAA | TGAAAAATAG | TCTCTAGTGA | AAGCGAAGAC | TTCGCGCACT | 27000 |
| 5 | GTAATTGCGA | CATATTTTGC | AACACACATT | TATTAAAATC | TGTTGTAAAA | CGTTGACATA | 27060 |
| | CATCTTTTGC | ATCGATŢGTT | AACGTTAACG | GCACAATATT | TCCGTGTAAA | TCATTTGGTA | 27120 |
| 10 | AATGTGATGG | TACATGTATA | CCTAATGTGA | CATCATGTTG | TTGACTCATT | ATATGATTAG | 27180 |
| | CTAAATACAC | ACTAACAGCC | AACGATGCCA | TATCTATTGA | TGTCATATCA | TCAATCAAAT | 27240 |
| | ACGTTTGATA | TAAAGCTTGT | TCAAATGGAT | GCTTAATTGG | AAAATAACTA | TCAATATGGA | 27300 |
| 15 | TGTCAGAGTT | ATTCTCTAAC | CGAAAATAGT | TTGAGTCTAA | TGCTATATGC | GATGCATCTT | 27360 |
| | GATTGTCTTT | ATCATCATTT | CTATTTATAT | GTGCATGCTG | TCGAGTATTG | TTAATAACAG | 27420 |
| | TATTGCCACG | ATATGCATTG | CATAAATCAT | CAAGAANAAT | ATCAATTTGA | CTATCATCga | 27480 |
| 20 | AAATGGaCAC | ATGAAAATCT | aATAGTATAT | ATGCAGCATC | AGCGAACTGm | AACAATTTAA | 27540 |
| | CTTTGAATAA | AGGTGAATCA | TTAAAATGGT | AAGTACTTAA | TTCTTGCTTA | AAAAAAGCTT | 27600 |
| | СТАААТСАТА | GTTTGCGGAA | GAAGATGGAA | CTTGTTTTAT | CTCAATAAAA | GGCAGAAATT | 27660 |
| 25 | CATGAAGTAT | CATTGTTAAA | TTGTCATCGG | TAGTAACATC | ÄAAAAAATGT | CTTATAGATG | 27720 |
| | CATGTTGTgC | ACAATTGTCG | ATAATGCATA | CATCATTTTA | GTAGCTTCAA | CATTTTTAGC | 27780 |
| | GAGTTTAACC | CAATACGCAT | TACGGTGTGT | CGTTGATTCT | GTATTATTTT | TGTATATACG | 27840 |
| 30 | AAAATATTCC | TGTTGAAATC | TCAAATTACC | CATAATCATA | AAAAGTCCTT | CTTTCATATC | 27900 |
| | ATAATACTCA | TTACTTACTG | AAATTGCATG | ATGATATGAT | AACCGACGAA | ATGTTĄATTA | 27960 |
| 35 | ACTCGTTATG | TAATGATTAA | TATAAAACAC | CATTCGCAAC | ATATGAGCGA | TATATTCTAC | 28020 |
| | CCTAAAATAC | ATCTTGTATC | ATCGTTACAA | TTGGTATATT | TTTCAATGTA | AATTACATAC | 28080 |
| | ATCTTCGATA | AATAGCACAC | TACAAATCGT | TAATCACTTT | CTGTTGTTCA | CATCTCATTG | 28140 |
| 40 | CAAACTCAAT | ATTGTTGTTA | CAAAATATCC | ATGAAGCAAG | TTTATATTAA | ACAAACAACT | 28200 |
| | CGCATAAAAC | AATTGTTATC | CTTAAATTTT | AACAAATTCT | TAATAAATTT | ATCTCTATTT | 28260 |
| | TAATTACGAC | CAAATTAATA | GGTTTTCCAT | ATAAAAAGAT | GCATAAAATA | AATATTTAAA | 28320 |
| 45 | TAAATTCAAT | TTGTATTCAC | TIGITITIGT | CCCCCAAATA | CACCAGCAAC | AAGCATGCTA | 28380 |
| | GCACCAATTG | TTAAAACGAT | AAACATATAC | AGTCCCATTI | GTAATGACGT | TAAGAAAACA | 28440 |
| | CCCAACACAA | TCCCTAACCT | AGCTAGTGTT | TCTGAAAAAT | GAATACCTAA | TGCATTAACT | 28500 |
| 50 | GCACTATATG | TICCTCTTT | AGCTTTAGGA | ATAATTTTAA | AGCGTTGTTC | TGAAACTATA | 28560 |
| | GGCGAATAAA | TAATTTCACC | TACAGTCGCA | ATTATCATAA | AAACAACTAA | TAAGCCAAAC | 28620 |

| | GCTTTTTTAA | AATCTATTT | ' CAATACAACT | TTCGAGATTG | AATACGTGAG | TAAAATGACG | 28740 |
|---|------------|------------|--------------|------------|------------|------------|-------|
| | ACGACCGTAT | TAATCATTAG | CAAGATTGCT | AACATCTTAG | CACCTGTAAT | ATCATATGAA | 28800 |
| | CCTATACTTA | TIGTTTCAAA | CTGATCCTTT | AGTCTAATAG | CAATATATGA | GGAGATTGAA | 28860 |
| | AATTCACCCA | TCATGATGAT | ACTGAACCCC | GAAATCAATA | ACATATAATT | ACGGTCTTTC | 28920 |
| | AAAACTAATT | TATAACTGCG | AAATATATTC | ATTATTTGTA | ATTTTTGATA | ACGACTTGCA | 28980 |
| | TGCCTCTTGT | CATCACTTTG | CTTTACTTGA | TTTCGGTCTT | GAGGTAACCA | AATATATAAA | 29040 |
| | ATAAAGAGTA | CAATTAAAAA | TATACAAGCT | GCTATTAAGA | AAAGTAGTAA | CATACTGTAG | 29100 |
| | CCATACATCA | AGCCACCTAA | CAATGCCCCA | ATAGCTACCG | ATAAGTTTGT | CATCCAATAG | 29160 |
| | CTAATCTTGT | AAATATAATG | TTCCACGTCT | TCGGTAATTG | CATCCATAAT | TAATGTGTCC | 29220 |
| | ATAACTGGAA | ATTGTAATCC | CCAAACGATT | GTAAATATGG | CATATGCAAC | ACAAAAACCA | 29280 |
| | ATAATTTGCC | ACAATTGATG | TGACCCAAAT | ACGCCCATGA | ACACAAGCAT | TATCACCATC | 29340 |
| | GTCGCTTGAT | AAATAAGTAC | TAGCAACTIT | tTCGGAAATA | TCTCAATAAG | GTAACCAGAT | 29400 |
| | ATAATGGACA | ATGGAAATTT | nAGAACCACT | AAACCAACAA | GATATATACC | GACAATTGAT | 29460 |
| | TGACTTAACA | TATCTGTTAA | ATATAGTGCT | ATAAACGGTA | TAAATGCTGT | CGTAATAATT | 29520 |
| • | AGCTGTAAAA | nATTGCTAAT | CAATCGTACT | TTCAA | • | | 29555 |

(2) INFORMATION FOR SEQ ID NO: 207:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1539 base pairs
- (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 207:

| AAAAAAAAA | AAAAANGGTG | AATCTTTAAT | TAAACACTAA | TATTGTAAAA | GATGTTAAGT | 60 |
|------------|------------|------------|------------|------------|--------------|-----|
| AAACGCTTAA | TGACACTTAT | TTTTTGAAAA | TAATAGTAAT | ATCATTTTGT | TAAATGAAAG | 120 |
| AATAAAGCTA | TAATmATTAT | AGAATAACTA | TTTAAAGGAG | ATTATAAACA | TGCCAATTAT | 180 |
| TACAGATGTT | TACGCTCGCG | AAGTCTTAGA | CTCTCGTGGT | AACCCAACTG | TTGAAGTAGA | 240 |
| AGTATTAACT | GAAAGTGGCG | CATTTGGTCG | TGCATTAGTA | CCATCAGGTG | CTTCAACTGG | 300 |
| TGAACACGAA | GCTGTTGAAT | TACGTGATGG | AGACAAATCA | CGTTATTTAG | GTAAAGGTGT | 360 |
| TACTAAAGCA | GTTGAAAACG | TTAATGAAAT | CATCGCACCA | GAAATTATTG | AAGGTGAATT | 420 |
| TTCAGTATTA | GATCAAGTAT | CTATTGATAA | AATGATGATC | GCATTAGACG | CTACTCCA A A | 400 |

| | AGCTGACTTA | TTAGGTCAAC | CACTITACAA | ATATTTAGGT | GGATTTAATG | GTAAGCAGTT | 600 |
|----|-------------|--|--|------------|------------|------------|------|
| | ACCAGTACCA | ATGATGAACA | TCGTTAATGG | TGGTTCTCAC | TCAGATGCTC | CAATTGCATT | 666 |
| 5 | CCAAGAATTC | ATGATTTTAC | CTGTAGGTGC | TACAACGTTC | AAAGAATCAT | TACGTTGGGG | 720 |
| | TACTGAAATT | TTCCACAACT | TAAAATCAAT | TTTAAGCAAA | CGTGGTTTAG | AAACTGCAGT | 786 |
| | AGGTGACGAA | GGTGGTTTCG | CTCCTAAATT | TGAAGGTACT | GAAGATGCTG | TTGAAACAAT | 840 |
| 0 | TATCCAAGCA | ATCGAAGCAG | CTGGTTACAA | ACCAGGTGAA | GAAGTATTCT | TAGGATTTGA | 900 |
| | CTGTGCATCA | TCAGAATTCT | ATGAAAATGG | TGTATATGAC | TACAGTAAGT | TCGAAGGCGA | 960 |
| _ | ACACGGTGCA | AAACGTACAG | CTGCAGAACA | AGTTGACTAC | TTAGAACAAT | TAGTAGACAA | 1020 |
| 5 | ATATCCTATC | ATTACAATTG | AAGACGGTAT | GGACGAAAAC | GACTGGGATG | GTTGGAAACA | 1080 |
| | ACTTACAGAA | CGTATCGGTG | ACCGTGTACA | ATTAGTAGGT | GACGATTTAT | TCGTAACAAA | 1140 |
| 0 | CACTGAAATT | TTAGCAAAAG | GTATTGAAAA | CGGAATTGGT | AACTCAATCT | TAATTAAAGT | 1200 |
| | TAACCAAATC | GGTACATTAA | CTGAAACATT | TGATGCAATC | GAAATGGCTC | AAAAAGCTGG | 1260 |
| | TTACACAGCA | GTAGTTTCTC | ACCGTTCAGg | aaacagaaga | TACAACAATT | GCTGATATTG | 1320 |
| 5 | CTGTTGCTAC | AAACGCTGGT | CAAATTAAAA | CTGGTTCATT | ATCACGTACT | GACCGTATTG | 1380 |
| | СТАВАТАСАА | TCAATTATTA | CGTATCGAGA | TGAATTATTT | GAAACTGCTA | AATATGACGG | 1440 |
| | TATCAAATCA | TTCTATAACT | TAGATAAATA | ATTTTCTnTA | TAATCAAATG | CTGACATAAT | 1500 |
| | TTTAGTTGAG | GATTATTATG | ACGGTATAAA | TAAATAAAG | | | 1539 |
| | (2) INFORMA | TION FOR SE | Q ID NO: 20 | 8: | | | |
| 25 | - (| QUENCE CHAR (A) LENGTH: (B) TYPE: nu (C) STRANDED (D) TOPOLOGY | 846 base pa cleic acid NESS: doubl | irs | | | |

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 208:

| CAATTTCTAT | CTATCAATGA | TGTGCATACT | TCCAnTTAAA | TTAAtCGAAA | TGaATCAAGG | 60 |
|------------|------------|------------|------------|--------------------|------------------------|-----|
| TATATCATTC | CTGCCTCTTT | ATATAACaAC | AAATAGTGAT | TACAATATTT | CGGTTATTAA | 120 |
| CACGAAAATT | TTACAAGCAC | CTATTTCATT | TACATATATA | TACAGCAAAA | AAGAAAGCCC | 180 |
| AGAAATATTG | GTGTTTATTA | AATCATTTAA | AAAGTATATT | GCCAATGAAC | AATTATAATA | 240 |
| AATTTCAAAT | CTAAAAAACC | AAGAATGCGA | TTAATCATCA | CATTCTTGGT | TCAATTTTAT | 300 |
| TCATGAATT | тттсаасатт | AAACGTTAAG | TTATTCTCTC | ידי ה ה הידידי ה ה | እ እ ርጥጥ የእ የነገር | 260 |

| | CGTTGTACAA | AACGTTTTAA | TGGTCTTGCA | CCGTATTGAG | GTTCATAAGC | TTCTTGACCT | 480 | | | |
|-------------------------------------|------------|------------|------------|------------|------------|------------|-----|--|--|--|
| | AGCCAAGCTT | TAGCATCATC | AGAAACTTCA | ATTGAGATTC | GTTGTTCTAA | TAATCTTATA | 540 | | | |
| | TTTAATTGCG | TTAAGATTTT | ATCTACAATC | ATACTCATGT | CATCAATAGA | TAATGGTTTA | 600 | | | |
| | AATAATACGA | TATCATCCAT | ACGATTCAAA | ATTTCTGGTT | TGAAATATGC | ATTTAAACTT | 660 | | | |
| | GTCATAACAG | CTTTTTCTGT | TGATTCTGTA | ATTTCACCAG | TCTCTTTTAC | GTTTTCTAAT | 720 | | | |
| | AAAACTTGAG | ATCCAATATT | ACTTGTCATA | ATAATAATAG | TATTTTTAAA | ATCAACGCTA | 780 | | | |
| | CGTCCTTTAG | AATCAGTTAA | ACGGCtTCAT | CTAAAATTTG | CAATAATACA | TTAAAGACGT | 840 | | | |
| | CAGTAT | | | | | | 846 | | | |
| (2) INFORMATION FOR SEQ ID NO: 209: | | | | | | | | | | |

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1674 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 209:

| nTGGGAACAG | TAAGCCAGTA | TTTTTAGAAA | GTTGCCATAC | ATGAGCGTCG | ATTTTTCCAA | 60 |
|------------|------------|------------|------------|------------|------------|-----|
| TATGGCTATG | ACACTAGAAC | AATGGGAATT | TGGAGGAAAA | GTAAATGATT | AAACCTAAAA | 120 |
| TAGCATTAAC | CATTGCAGGT | ACTGATCCAA | CAGGTGGTGC | CGGCGTAATG | GCTGATTTAA | 180 |
| AATCATTTCA | TTCATGTGGT | GTATATGGTA | TGGGCGTCGT | TACAAGTmTT | GTTGCTCAAA | 240 |
| ATACATTGGG | CGTACAACAT | ATTCATAATT | TAAATCATCA | ATGGGTAGAT | GAACAACTTG | 300 |
| ATAGTGTCTT | CAATGATACC | TTACCTCATG | CTATTAAAAC | GGGGATGATT | GCTACAGCAG | 360 |
| ATACTATGGA | AACGATTCGT | CATTATTTAA | TGCAACATGA | ATCTATTCCA | TATGTAATLG | 420 |
| ATCCTGTTAT | GTTGGCGAAA | rCggTGATTC | ATaATGGwTA | ATGACACAAg | CaAAACTTGC | 480 |
| AGCATACGTT | ATTGCCATTA | GCTGACGTAG | TAACACCGAA | TTTACCAGAA | GCTGAAGAAA | 540 |
| TAACGGGACT | AACCATTGAT | AGTGAAGAAA | AAATTATGCA | GCTGGCCGC | ATCTTTATTA | 600 |
| ATGAGATTGG | TAGTAAAGGT | GTCATCATTA | AAGGCGGTCA | TTCAAATGAT | ACTGATATAG | 660 |
| CAAAAGATTA | TTTATTTACT | AACGAAGGTG | TTCAAACATT | TGAAAATGAA | CGATTTAAAA | 720 |
| CAAnACATAC | GCATGGAACA | GGGTGTACAT | TTTCAGCAGT | TATAACGGCA | GAACTTGCAA | 780 |
| AAGGTAGACC | ATTATTTGAG | GCTGTACACA | AGGCTAAAAA | GTTTATTTCA | ATGAGTATAC | 840 |
| AATATACGCC | TGAAATCGGC | CGTGGTAGAG | GTCCAGTGAA | TCATTTTGCA | TATTTAAAGA | 900 |

55

5

10

15

20

30

35

40

45

| TGTTATACAA | ACGATGTAGT | TAAAAATTTT | ACAGCGAATG | GTTTATTAAG | TATTGGTGCT | 1020 |
|-------------|-------------|-------------|------------|------------|------------|-------|
| AGCCCTGCAA | TGAGTGAAGC | TCCCGAAGAA | GCTGAAGAAT | TTTACAAAGT | TGCACAAGCG | 1080 |
| CTATTAATCA | ATATCGGTAC | TTTAACAGCA | GAAAATGAAC | AAGATATTAT | TGCGATTGCT | 1140 |
| CAAACGGCAA | ATGAGGCAGG | CTTACCTATT | GTATTTGACC | CTGTAGCTGT | TGGTGCTTCT | 1,200 |
| ACATATCGAA | AGCAATTTTG | TAAATTATTA | TTGAAATCAG | CGAAAGTATC | AGTAATTAAA | 1260 |
| GGCAATGCAT | CTGAAATATT | AGCGTTGATT | GATGATACAG | CAACTATGAA | AGGTACAGAT | 1320 |
| AGTGATGCTA | ATCTTGATGC | GGTTGCAATA | GCGAAAAAGG | tTACGCAACA | TATAAAACTG | 1380 |
| CAATAGTAAT | CACAGGTAAA | GAGGACGTTA | TTGtTCmAGA | TAATAAAGCC | TTCGTATTAG | 1440 |
| CTAATGGATC | TCCATTATTA | GCACGAGTAA | CTGGAGCTGG | TTGTTTATTA | GGAGGCGTTA | 1500 |
| TTGCTGGATT | TTTATTTAGA | GAAACAGAAC | CAGACATAGA | AGCGTTAATT | GAAGCGGTAA | 1560 |
| GCgkATTTAA | TATTGCTGCT | GAGGTAGCTG | CTGAAAATGA | AAATTGTGGT | GGTCCTGGTA | 1620 |
| CGTTTTCACC | ATTGTTGCTT | GATACGTTAT | ATCATTTAAA | TGAAACAACC | TATC | 1674 |
| (2) INFORMA | TION FOR SE | Q ID NO: 21 | .o: | | | |

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2232 base pairs

 - (B) TYPE: nucleic acid(C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 210:

| A. | IGAGTTGCC | GATGAATTTA | GCACCACCAA | CGATTGCnTT | TGATACTGTG | TCCCAACCAG | 60 |
|----|-----------|------------|------------|------------|------------|------------|-----|
| C. | ITGTTTAGC | ATATTTAATA | CCTTCACGTA | AAGGATCGTT | ATCATATGCA | GCAATACCAA | 120 |
| A. | TACGTTATG | GTATTTCGTG | TTTGAGTTAG | TTACAACTTT | GTnTTGCACT | ACATCTGCAC | 180 |
| C. | ITTCGCTAA | TTGAGAAGTA | CCGTTACCTG | TTTCTAATAG | GGCATGTGAG | ATAAGATAAA | 240 |
| C. | ITCATTAAT | GCCATACATT | TGAGCAGCTT | TGTTAAATGC | AGCACCTTGG | TTTTCTAATA | 300 |
| C | ACCTTTACC | TTTTAAGAAT | TGATTAATTT | TATCAATAGA | AATATTTTGT | GGTTGGTCTA | 360 |
| A | GCGTAAGAA | TTGATATTTT | AATGCTGGAT | CTTGAGCTAA | ACGCTTCGTA | TCCATTGCAT | 420 |
| G | CTTAACATC | ATTAAATTTA | GCATCTGTCC | ACTTACCTGG | TACACGTTGT | ACTTGTGGTT | 480 |
| T | ATATTGTAA | ACCAGCTTGT | ATTTGAGCAA | CTTGGTTTAA | TGTCATACCT | GTTTGATTAT | 540 |
| A | CTTAATTAA | TTCTTTAGCT | AAATCAGTTG | ATTTAATCCA | TGCtAATTTA | CCGTTAGATA | 600 |
| A? | TTACCATA | GTACCAAGTT | TGTCCATTAA | TGACTTGTTC | TTTAACAACT | GCGAATGGTT | 660 |

| | AACCATTACC | ATTTTTAATT | ACATAAGTGT | AGTTATAATC | TTTGGcAGCT | GATGTAGTTG | 780 |
|----|------------|------------|------------|-------------|------------|------------|------|
| 5 | GTTTCACAGC | AGTTGGTGCA | GTTAAATCTT | TTGCATTTAC | CCAACCAGTG | CGGTTATTAA | 840 |
| | TAGTACCGTA | TAAATAAACA | TCTTTGCCTA | CAGATACTTG | TTTCGTTGCA | TTAAATGTAC | 900 |
| | CTTGAGCAAT | GTTATTGCCT | GTTAAAATGA | CTTGGTTTTT | AGTACCCCAA | GGAACCATTG | 960 |
| 10 | ATAAGCCGTT | ATTTGATTTA | TTAACAGTAT | ATTTTTGAGT | CGTTTTAACT | TCTTTGCCTA | 1020 |
| | AGTTTTGAAC | ATTTAAGTCT | TTTACATTGA | ACCAACCTAA | TGGGATGTTA | TGGCTTGTAT | 1080 |
| | TGTTTAATAA | TACATACGTT | TCATTACCAT | GAGCACGCTC | TTTTGTTACA | TAGAACGTAC | 1140 |
| 15 | GGTCTGCATA | TTTCGCACCG | TTTTTCGCTG | TTTTTTCATA | AACAGAAGCA | CGAATACCAG | 1200 |
| | TGTTGTTTGG | TTTAACTTGA | GCAATCTTGC | TAACTGTTTG | AGTCGTTTGT | GGTTTAGTAA | 1260 |
| 20 | CAGTATAAGC | TTTTACAGCT | GTTTTTGGTT | GTGCTACTGC | TTTTTTAGGT | GCAGCAGGTA | 1320 |
| | CAGCTAAATA | TGCTTTACTT | ACCCAACCAG | ATTTACCATT | TACAGTTCCA | AATAAATAGA | 1380 |
| | TAGATTTATC | AATTTGTTGT | TGCTTAGTCG | CTTTAAAAGT. | TTGGTTACCT | GTACCAGAAA | 1440 |
| 25 | CTGCACCAGC | TTCTTGTTTA | TAAGTGCCCC | AAGGTACTGA | ATATAATTTA | GTGCCTGGgT | 1500 |
| | TTACTGTATA | TGTTTGCATT | ACATTTACAG | GTGATTTTGC | ATtGtTATAA | ATACGTCACC | 1560 |
| | TTGTTTAACC | CAACCAATTA | AAGTTGGACT | ATTGTAATCT | TTAACTAAGT | AGAATTTGTT | 1620 |
| 30 | TCCACCTAAA | CTTGCTTCTT | TTGTTACAGC | AAATGTTTTT | TGAACTTCTT | TCGTTGGCTT | 1680 |
| | ACCAGTTTTG | TCATAAACTG | TAGTGAATAA | GCCATTGTTT | TTAGCATTAA | TTTGAGCAAC | 1740 |
| 35 | ACCGTTTAAT | GATGAAACTG | TTAATTTATT | ATTTGTTGTA | GGTGTTGATG | GCTTAGGTGT | 1800 |
| 33 | TGGTGTAGGC | GTAGGTTTAG | CAGTATCAAC | TAAATATGCT | TTACTTACCC | AACCAGATTT | 1860 |
| | ACCATTCACA | GAGCCATATA | AATAAATTGA | TTTATCAATT | TGTTGTTGCT | TTGAAGCCTT | 1920 |
| 40 | AAAŢGTTTGG | TTTCCAGAGC | CAGACACACT | ACCAGCAACT | TGTTTAGATG | TACCCCAAGG | 1980 |
| | TACTGTATAA | AGTTTCGTAC | CAGGTTTGAT | TGAATATGAT | TGATTTACAT | TTACAGGTGA | 2640 |
| | TTTAGCTGTG | TTGTAAACCA | CATCGCCTTC | TTTAACCCAA | CCAAATTTAT | TACCAGAATT | 2100 |
| 45 | GTAATCTTGA | ACAAGATAGA | ATTTTTGATT | ACCTAATGTA | GCTGTTTTAG | ATACAGCAAA | 2160 |
| | TGTTTTTTGA | ACTTCATTAG | TTGCTTTACC | AGTTTTGTCG | TATACAGTAG | TATATAAACC | 2220 |
| 50 | ACTATTTGTT | GG | | | | | 2232 |
| 50 | | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 211:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2082 base pairs (B) TYPE: nucleic acid

 - (C) STRANDEDNESS: double

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 211:

| 5 | GATTTAAATA | AAATTAATGG | ATATCGTGAT | CGTACGATGT | TAGAACTTCT | GTACGCAACG | 60 |
|----|------------|------------|------------|------------|-------------|------------|-------|
| 3 | GGAATGCGTG | TATCTGAATT | GATACATTTA | GAGTTAGAAA | ACGTGAACTT | AATAATGGGA | 120 |
| | TTTGTACGCG | TATTTGGTAA | AGGCGATAAA | GAAAGAATTG | TACCATTAGG | CGACGCAGTC | 180 |
| 10 | ATTGAGTACT | TAACTACTTA | TATTGAAACG | ATTAGACCGC | AACTTTTAAA | AAAGACTGTT | 240 |
| | ACTGAAGTCT | TATTTTTAAA | TATGCATGGT | AAACCTTTAT | CACGACAAGC | AATATGGAAA | 300 |
| | ATGATTAAAC | AAAATGGTGT | AAAGGCAAAC | ATTAAAAAGA | CGTTAACGCC | ACATACGTTA | 360 |
| 15 | CGCCACTCTT | TTGCGACACA | TTTATTGGAA | AATGGCGCAG | ATTTAAGAGC | AGTGCAAGAG | 420 |
| | ATGTTAGGTC | ACTCTGaCmT | ATCTACTACC | CmaCTCTATA | CmCATGTTTC | Graatctcaa | 480 |
| 20 | ATTAGAAAAA | TGTATAACCA | ATTTCATCCT | AGAGCATAAA | GTGAACAATA | ACTCAAAAGT | 540 |
| 20 | CACAATACAC | ATGACTAAAA | ATGTCTGTGC | TATTGTGGCT | TTTTTAAATT | GGTTGATTAA | 600 |
| | TTACGTCTAT | GTTTTCTTAA | TTGAATCGCT | TCTTCTTTTG | CTGCAATCAC | TTCTGAACGA | 660 |
| 25 | TCACGGCGCA | TGTGATGGTC | TACAATAAAA | GGATCTGTTG | CTGTTTCCTG | ATTATAATCA | 720 |
| | TAGTCTGGAT | AGTTGGCCTT | GATGATGCGT | TCAAAGACTG | GAGTTATTGG | TAATATAACA | 780 |
| | GATGAAAAAG | GCTTTGCTGC | ATTCAATTTT | GCAATCTGTT | GCTCAATTAA | CAACTGATAA | 840 |
| 30 | TCATTTAAAT | TAAGGTATAA | CGCATCTCTA | TCTTTAGCAT | TTTGTATTAT | TTCTTTAGAT | 900 |
| | TTATTAAAAG | ACTTATAGGC | GCCTTTTAAA | TTATTGCGGC | GATAATGGTA | ACAAGCAGTT | 960 |
| | GCAAACAAGA | TTAAACTAAC | AACTGCATCT | TGCTTACTGT | AGTTATTTTC | AGCTTTCCAT | 1020 |
| 35 | GCATCTTCTA | AAATGTCATG | ACATAGGAAA | TAATGTTGCT | TAGTATGAAA | TTGATAATAG | 1080 |
| | AAATTTATCA | GTGCCTGTTG | CATTTTGTTA | TCACCCCAAT | TTAAAAGTAA | GTTATTTTCA | 1140 |
| 40 | TGCTATAATA | TTTTAGAGAA | TTATGCACAT | ATGACGCAAT | ACGAGGTAGA | TATTATGTAT | 1200 |
| | GAAGTTAAAT | TAGATGCTTT | CAATGGACCA | TTAGATTTAT | TGCTGCATCT | TATCCAAAAA | 1260 |
| | TTTGAAATAG | ATATTTATGA | TATTCCTATG | CAAGCATTAA | CAGAGCAGTA | TATGCAGTAC | 1320 |
| 45 | GTTCATGCAA | TGAAACAGCT | TGAAATTAAT | ATTGCAAGTG | AATACCTAGT | ATTAGCGTCA | 1380 |
| | GAACTCTTAA | TGATTAAAAG | TAAGATGCTA | TTACCACAAT | CAACATCAGA | TATGGATGTT | 1440 |
| | GATGATGACC | CACGGGAAGA | TTTAGTtGGG | CGTTTAATAG | rATATCAAAA | TTATATAGAA | 1500 |
| 50 | TATACTGCLA | TTTTAAATGA | CATGAAAGAA | GAAAGAGATT | TTTATTTTAC | CAAAAAGACC | 1560 |
| | CACACATTTA | TO-CATION | *** | ATC | CC- NATCATE | 0030003000 | 1.636 |

| ATCTGTTGAA | ATCCGAAAAG | AGACATTTAC | CATTCAACAA | GCTACAGAAC | AAGTGACATC | 1740 |
|------------|------------|------------|------------|------------|------------|------|
| GAGATTGAAA | GATAAAGATC | ATTTTAACTT | CTTTAGTCTG | TTTACGTTTT | CTGAGCCAAT | 1800 |
| TGAACAAGTA | GTCACTCACT | TTTTAGCTAT | TTTAGAGATG | TCAAAAGCAG | GAATAATTAA | 1860 |
| TATTGAGCAA | CAACGTAATT | TTGAAGATAT | TAACATTATT | AGAGGAGTGA | ACTACCATTT | 1920 |
| TGGATAATCA | TGGTATATTA | GAGTCGCTTT | TATTTACAGC | TGGCGATGAA | GGTTTAGATG | 1980 |
| AAAAACAACT | ATTAGAAATA | TTAGATATGT | CGAAAGACCA | ACTCGTTGAA | TTAATTGAAA | 2040 |
| ATTATTCATC | ACATGGATTA | ATGATACAAC | GATTTGGAAT | GA | | 2082 |
| | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 212:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4219 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 212:

TCTATTCTCG TTCTTCCAAG ACCCTGAATT AGAAGTTAAG AAAATCGAAG AAGATGAGAA 60 AGAATCTATT AAAAAAGCTC AAAAAGGTAT TTATAAAGAC CCTAGAGACA TCAATGATGA 120 CGAACAAGAT GATGATACAA AAGATACTGT TGATAAAAAG GAATGATTGT AATTGCCTAA 180 CAAAAACACT CAAGAATATT GGGAAGAACG CGGACGCAAA GCAATCGAGA ATGAGTTGAA 240 GCGTGATAAA ACTAAAGCTG AAGAAATAGA ACGTATATTG AATATGATGA TTAAGCGCAT 300 360 AGAAGCACAA AAGATTATTG ATGAGTTCGA TGTAAAAGCG TTTCAAGAAG AAGCAAAAAG 420 ATTGGTCGAA AACAAGGAGT TTAGCGATAG AGCAAATGAA GAATTAAAGA AGTATAACAC 480 GAAAATGTAT GTATCTAGAG AACAGATGTT AAAGATTCAA ATAGAATTCT TAATTGCTTA 540 TGCAACAGCT CAAACAGAAT TATCGATGAG GGAATATTTC GAATCAACAG CTTATCGTGT 600 GTTCAGTGAT CAAGCGGGTA TTTTAGGTGA AGGTGTACAA GTAGCTAAAG AAGTTATAGA 660 TACAATCGTT GATACACAAT TTCATGGTGT CGTTTGGTCA GAGCGATTAT GGACTAATAC 720 CGAAGCAATG AAACAAGAAG TAGAAGAAAT AATTGCTAAT GTAGTTATTA GAGGTCGACA 780 TCCTAATGAA TATGTTAAAG ATATGCGCAA CACTTAAATA AATTCGAAGG CACAGCACGA 840 CAAAAGACCG CAGCAATTAA ATCATTGCTT TATACGGAAT CGGCACGTGT TCACGCACAA 900 TCAAGCATTG ACAGCATGAA AGAAATTTCA CCGGAAGGAT ATTATATGTA TATTGCAAAA 960

5

10

15

20

25

30

35

40

45

50

| | GACGCTAAAA | TTGGTGTTAA | TTTCTATCCT | ATGCATATCA | ATTGTCGTTC | AGATTGCGCT | 1080 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | TTACTACCTA | AATCTATGTG | GCCGAAAAAA | CCAAGCAAGA | AACGAAAAAC | AAAATACTTC | 1140 |
| 5 | GGAGGGAAAG | TGAAAAGCGG | TGATTGATTT | AAAAGTGAAG | TTTTTTAAAG | GCAAGTTAGT | 1200 |
| | TTTGTATGAC | AGTAAATTAA | ATGTTTGGAG | GATACTAATA | TGAGTAATAC | TGACAAATAC | 1260 |
| 10 | CTTAGAGACA | TAGCAAGAGA | ATTAAAAGGT | AŢACGTAAAG | AGTTACAAAA | GCGAAACGAA | 1320 |
| 10 | ACAGTTATTA | TTGATGCAAA | CTTAGACAGT | TTAAGGTCGG | CAGTATTAGC | CGATAAAGAA | 1380 |
| | AAATCGAAAT | ATAATGAACC | TCTCTTTTAA | TAGCTAGCAC | TTAATTGTGT | TGGCTATTTT | 1440 |
| 15 | TTATGTCCAA | AACGTGCTGA | TGACATAAAA | AGCACGCATG | GAAAAACAGT | CGACAGACTA | 1500 |
| | TAAATGGAGG | TATATCTCAT | GGAAGAAAT | AAACTTAAGT | TTAATTTGCA | aTTTTTTGCA | 1560 |
| | GACCAATCAG | ATGATCCGGA | CGAACCAGGC | GGAGATGGTA | AAAAAGGAAA | TCCTGATAAG | 1620 |
| 20 | AAAGAAAATG | ACGAAGGTAC | TGAAATAACT | TTCACGCCAG | AGCAACAAAA | GAAAGTTGAT | 1680 |
| | GAAATACTTG | AACGTCGTGT | AGCCCACGAA | AAGAAAAAG | CTGATGAGTA | TGCAAAAGAA | 1740 |
| | AAAGCAGCAG | AAGCTGCTAA | AGAAGCTGCT | AAATTAGCGA | AAATGAACAA | GGATCAAAAA | 1800 |
| 25 | GATGAATATG | AACGCGAACA | AATGGAAAAA | GAACTGGAAC | AATTACGTTC | AGAAAACAA | 1860 |
| | TTAAACGAAA | TGCGTTCAGA | AGCACGAAAA | ATGTTGAGTG | AAGCGGaAGT | TGATTCATCA | 1920 |
| 3 <i>0</i> | GATGrGGTTG | TCAATTTAGT | TGTAACAGAT | ACTGCTGAAC | AAACTAAATT | GAATGTTGAA | 1980 |
| | GCTTTTTCTA | ATGCAGTAAA | AAAAGCGGTT | AATGAAGCGG | TTAAGGTTAA | CGCTAGACAA | 2040 |
| | TCGCCATTGA | CTGGTGGAGA | TTCATTTAAT | CACTCGACTA | AAAATAAACC | GCAAAACTTA | 2100 |
| 35 | GCTGAAATAG | CTAGACAAAA | AaGAATTATT | AAAAATTAAC | GGAGGCATTT | AAATGGAACA | 2160 |
| | AACACAAAAA | TTAAAATTAA | ATTTGCAACA | TTTTGCAAGT | AACAATGTTA | AACCACAAGT | 2220 |
| | ATTTAACCCT | GACAATGTAA | TGATGCATGA | AAAGAAAGAT | GGCACGTTGT | TAAACGACTT | 2280 |
| 40 | TACAACACCT | ATCTTACAAG | AGGTTATGGA | AAACTCTAAA | ATCATGCAAT | TAGGTAAGTA | 2340 |
| | CGAACCAATG | GAAGGTACTG | AGAAGAAGTT | TACTTTTTGG | GCTGATAAAC | CAGGTGCTTA | 2400 |
| 45 | CTGGGTAGGT | GAAGGTCAAA | AAATCGAAAC | GTCTAAGGCT | ACTTGGGTTA | ATGCTACAAT | 2460 |
| 4 3 | GAGAGCGTTT | AAATTAGGGG | TTATCTTACC | AGTAACAAAA | GAATTCTTGA | ATTACACTTA | 2520 |
| | TTCACAATTC | TTTGAAGAAA | TGAAACCTAT | GATTGCTGAA | GCTTTCTATA | AAAAGTTTGA | 2580 |
| 50 | CGAGGCAGGT | ATTTTGAATC | AAGGTAACAA | TCCGTTCGGT | AAATCAATTG | CACAATCAAT | 2640 |
| | TGAAAAAACT | AATAAGGTTA | TTAAAGGTGA | CTTCACACAA | GATAACATTA | TTGATTTAGA | 2700 |
| | GGCATTGCTT | GAAGATGACG | AATTAGAAGC | AAATGCATTT | ATCTCAAAAA | CACAAAACAG | 2760 |

| | TGATTCGTTA | GACGGTCTAC | CTGTGGTTAA | CCTTAAATCA | AGCAACTTAA | AACGTGGTGA | 2880 |
|----|------------|------------|------------|------------|------------|------------|------|
| | ATTAATCACT | GGTGACTTCG | ACAAATTGAT | TTATGGTATC | CCTCAATTAA | TCGAATACAA | 2940 |
| 5 | AATCGATGAA | ACTGCACAAT | TATCTACAGT | TAAAAACGAA | GATGGCACAC | CTGTAAACTT | 3000 |
| | GTTTGAACAA | GACATGGTGG | CATTACGTGC | AACTATGCAT | GTAGCATTGC | ATATTGCTGA | 3060 |
| 10 | TGATAAAGCG | TTTGCTAAGT | TAGTTCCTGC | TGACAAAAGA | ACAGATTCAG | TTCCAGGAGA | 3120 |
| | AGTTTAATAA | ATAATTAGGA | GTGGTAACAT | GCCCGAAATC | ATTGGAATTG | TTAAAGTAGA | 3180 |
| | TTTTACAGAT | TTAGAAGATA | ACAGACATGT | CTATATGAAA | GGGCATGTCT | ACCCTCGTAA | 3240 |
| 15 | AGGTTATAAT | CCTACAGATG | AACGTATCAA | AGCTTTAGCT | AGTGTTGAAA | ATAAACGCAA | 3300 |
| | CAAACAAATG | ATTTACATTG | TAAATGACAA | ATTAACCAAA | AAAGAACTTG | TCGAAATAGC | 3360 |
| | AAGTGTTGCT | GGCTTACAAG | TTGATGAAAA | ACAAACAAAA | GCTGAAATTA | TCAATGCTTT | 3420 |
| 20 | TGAGTCACTA | GAGTAGGTGG | TTATATGACT | ACGCTAGCTG | ATGTAAAAAA | ACGTATTGGT | 3480 |
| | CTTAAAGATG | AAAAGCAAGA | TGAACAATTA | GAAGAAATCA | TAAAAAGTTG | TGAAAGCCAG | 3540 |
| 25 | TTGTTATCAA | TGTTACCTAT | TGAAGTTGAA | CAAATACCGG | AAAGgTTTAG | TTACATGATT | 3600 |
| 23 | AAAGAAGTTG | CAGTTAAACG | CTACAACAGG | ATTGGTGCTG | AAGtATGACA | TCAGAAGCGG | 3660 |
| | TTGACGGACG | TAGCAATGCG | TATGAATTGA | ACGATTLCAA | GGAGTATGAA | GCTATTATTG | 3720 |
| 30 | ATAATTACTT | TAATGCTAGA | ACGAGAACTA | AAAAAGGAAG | GGCTGTGTTC | TTTTGAGATA | 3780 |
| | TGAAGATAGA | GTTATTTTTC | AATTAGAACA | AGTAGCAACT | TACAATCCTA | AAACTAGCAA | 3840 |
| | AAAAGAAAAC | ACACTAATCA | CTTATGATGC | GATACCATGC | AATATTAACC | CCATTTCTAG | 3900 |
| 35 | AGCAAGAAAG | CAACTTGAAT | TTGGTGATGT | AAAAAACGAT | GTAAGTGTTC | TGAGGATAAA | 3960 |
| | AGAATCAATA | | | | | | 4020 |
| | TGATACAAGG | ATATACAGAC | ACGAAACGTC | ATATTATATC | GAAGAGGTCA | ATTGATGAAT | 4080 |
| 40 | | | | | TGAAAACCAA | | 4140 |
| | GATGTAGATG | ATATTTTACA | GGAAAACGCC | AAAGAATATG | TAGTACGAGC | TAAATTGAAA | 4200 |
| 45 | GCTAGAGAAG | TAATGAATA | | Ÿ | | | 4219 |

(2) INFORMATION FOR SEQ ID NO: 213:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1999 base pairs
- (B) TYPE: nucleic acid (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

55

50 ·

| | GCTTACAAGT | ATATTCATAA | TTACATATTC | AAGGTCCTTG | CATGTGGTAT | TTTGCTATGG | 60 |
|----|------------|------------|------------|------------|------------|------------|------|
| | yCtTTaACTA | CAACGGGGTC | TAAGACTGCG | TTTATCATAT | TAATCGTCTT | AGCCATTLAT | 120 |
| 5 | TYCTTTATKA | AAAAGTTATT | TAGTAGAAAT | GCGGTAAGTG | TTGTGAGTAT | GTCAGTGATT | 180 |
| | ATGCTGATAT | TACTTTGTTT | TACCTTTTAT | AATATCAACT | ACTATTTATT | CCAATTAAGC | 240 |
| •• | GACCTTGATG | CCTTACCGTC | ATTAGATCGA | ATGGCGTCTA | TTTTTGAAGA | GGGCTTTGCA | 300 |
| 10 | TCATTAAATG | ATAGTGGGTC | TGAGCGAAGT | GTTGTATGGA | TAAATGCCAT | TTCAGTAATT | 360 |
| | AAATATACAC | TAGGTTTTGG | TGTCGGATTA | GTGGATTATG | TACATATTGG | CTCGCAAATT | 420 |
| 15 | AATGGTATTT | TACTTGTTGC | CCATAATACA | TATTTGCAGA | TCTTTGCGGA | ATGGGGCATT | 480 |
| | TTATTCGGTG | CATTATTTAT | CATATTTATG | CTTTATTTAC | TGTTTGAATT | ATTTAGATTT | 540 |
| | AACATTTCTG | GGAAAAATGT | AACAGCAATT | GTTGTAATGT | TGACGATGCT | GATTTACTTT | 600 |
| 20 | TTAACAGTAT | CATTTAATAA | CTCAAGATAT | GTCGCTTTTA | TTTTAGGAAT | TATCGTCTTT | 660 |
| | ATTGTTCAAT | ATGAAAAGAT | GGAAAGGGAT | CGTAATGAAG | AGTGATTCAC | TAAAAGAAAA | 720 |
| | TATTATTTAT | CAAGGGCTAT | ACCAATTGAT | TAGAACGATG | ACACCACTGA | TTACAATACC | 780 |
| 25 | CATTATTTCA | CGTGCATTTG | GTCCCAGTGG | TGTGGGTATT | GTTTCATTTT | CTTTCAATAT | 840 |
| | CGTGCAATAC | TTTTTGATGA | TTGCAAGTGT | TGGCGTTCAG | TTATATTTTA | ATAGAGTTAT | 900 |
| 30 | CGCGAAGTCC | GTTAACGACA | AACGGCAATT | GTCACAGCAG | TTTTGGGATA | TCTTTGTCAG | 960 |
| | TAAATTATTT | TTAGCGTTAA | CAGTTTTTGC | GATGTATATG | GTCGTAATTA | CTATATTTAT | 1020 |
| | TGATGATTAC | TATCTTATTT | TCCTACTACA | AGGAATCTAT | ATTATAGGTG | CAGCACTCGA | 1080 |
| 35 | TATTTCATGG | TTTTATGCTG | GAACTGAAAA | GTTTAAAATT | CCTAGCCTCA | GTAATATTGT | 1140 |
| | TGCGTCTGGT | ATTGTATTAA | GTGTAGTTGT | TATTTTTGTC | AAAGATCAAT | CAGATTTATC | 1200 |
| | ATTGTATGTA | TTTACTATTG | CTATTGTGAC | GGTATTAAAC | CAATTACCTT | TGTTTATCTA | 1260 |
| 40 | TTTAAAACGA | TACATTAGCT | TTGTTTCGGT | TAATTGGATA | CACGTCTGGC | AATTGTTTCG | 1320 |
| | TTCGTCATTt | AGCATACTTA | TTACCAAATG | GACAGCTCAA | CTTATATACT | AGTATTTCTT | 1380 |
| | GCGTTGTTCT | TGGTTTAGTA | GGTACATACC | AACAAGTTGG | TATCTTTTCT | AACGCATTTA | 1440 |
| 45 | ATATTTTAAC | GGTCGCAATC | ATAATGATTA | ATACATTTGA | TCTTGTAATG | ATTCCGCGTA | 1500 |
| | TTACCAAAAT | GTCTATCCAG | CAATCACATA | GTTTAACTAA | AACGTTAGCT | AATAATATGA | 1560 |
| 50 | ATATTCAATT | GATATTAaCA | ATACCTATGG | TCTTTgGTTT | AATTGCaATT | ATGCCATCAT | 1620 |
| | TTTATTTATG | GTTCLTTGGT | GAGGAATTCG | CATCAACTGT | CCCATTGATG | ACCATTTTAG | 1680 |
| | CGATACTTGT | ATTAATCATT | CCTTTAAATA | tGTTGaTAAq | CaGGCAATAT | TTATTAATAG | 1740 |

| TGGTCAATCA | TTATTTGCC | | | | | 1999 |
|------------|------------|------------|------------|------------|------------|------|
| ATATTGTAAG | TACGATTCAA | TGTGTCATTG | CTGCTGTTAT | GATGTTTATT | GTGCTTGGTG | 1980 |
| CAGAGTTTTT | CTTGCTCATT | TGGCGATTTA | TTGATATTAC | TAAAATCAAT | GTGAAGTTGA | 1920 |
| TATGTAYTAT | TTTGATATAT | TTTTATGGAA | TTTACGGTGC | TGCTATTGCG | CGTTTAATTA | 1860 |

(2) INFORMATION FOR SEQ ID NO: 214:

10

5

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7769 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

15

20

25

30

35

40

45

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 214:

| TCATTATTAA | GACTATTATA | TATAATGAAT | TTTAACTGGT | TTATTAAACG | AGAACGTCGG | 60 |
|------------|------------|------------|------------|------------|------------|------|
| GAATTAAGTA | ACTACAATAA | AAATAAGATA | TGACAATAAG | GAGACTACAC | GCGTGATCAT | 120 |
| TGCCATAATT | ATATTGATAT | TTATTTCGTT | TTTCTTTTCA | GGAAGCGAGA | CGGCATTAAC | 180 |
| GGCTGCCAAT | AAAACAAAAT | TTAAAACTGA | AGCTGACAAA | GGTGATAAAA | AAGCAAAAGG | 240 |
| CATTGTAAAG | TTACTTGAAA | AACCAAGTGA | GTTTATTACA | ACGATTCTAA | TTGGGAATAA | 300 |
| TGTCGCGAAT | ATTTTATTAC | CAACACTTGT | TACAATTATG | GCTTTACGTT | GGGGGATTAG | 360 |
| CGTTGGTATT | GCATCAGCTG | TTTTAACAGT | TGTTATCATT | TTGATCTCCG | AAGTGATTCC | 420 |
| CAAGTCTGTC | GCTGCAACAT | TTCCAGATAA | AATAACAAGG | CTTGTATATC | CAATTATTAA | 480 |
| TATTTGTGTC | ATTGTGTTCC | GTCCTATCAC | ATTACTTTTA | AATAAGTTGA | CGGACAGTAT | 540 |
| TAATCGAAGT | TTATCTAAGG | GCCAACCTCA | AGAACATCAA | TTTTCAAAAG | AAGAATTTAA | 600 |
| AACÁATGTTA | GCAATTGCTG | GACATGAAGG | TGCTTTAAAT | GAAATTGAGA | CGAGTAGGTT | 660 |
| GGAAGGTGTC | ATTAATTTTG | AAAATTTAAA | AGTAAAAGAT | GTAGATACAA | CACCTAGAAT | 720 |
| TAATGTGACG | GCATTTGCTT | CAAATGCGaC | ATACGAAGAA | GTTTATGAAA | CGGTTATGAA | 780 |
| TAAGCCATAC | ACTAGATATC | CAGTGTACGA | GGGAGATATT | GATAACATTA | TTGGGGTGTT | 840 |
| TCATTCTAAA | TATCTGTTGG | CTTGGAGTAA | TAAAAAAGAA | AATCAAATTA | CAAACTATTC | 900 |
| AGCTAAGCCA | TTATTTGTGA | ATGAACACAA | TAAAGCTGAA | TGGGTATTAC | GTAAGATGAC | 960 |
| TATTTCTAGA | AAACATTTAG | CAATTGTGTT | GGACGAATTT | GGTGGTACTG | AAGCGATAGT | 1020 |
| GTCACATGAA | GACTTAATTG | AAGAATTATT | AGGTATGGAA | ATTGAAGATG | AGATGGATAA | 1080 |
| AAAGGAAAAA | GAAAAACTTT | CTCAACAGCA | AATTCAATTT | CAACAACGGA | AAAATCGCAA | 1140 |

| | GTATTGAATA | TCCAATTATA | CAAGCAGGTA | TGGCAGGAAG | TACGACACCG | AAATTAGTTG | 1260 |
|----|------------|------------|------------|------------|------------|------------|------|
| | CATCAGTAAG | TAACAGTGGT | GGGTTAGGCA | CAATAGGCGC | AGGTTACTTT | AATACGCAGC | 1320 |
| 5 | aattggaaga | TGAAATAGAT | TATGTACGCC | AATTAACGTC | AAATTCTTTT | GGCGTAAATG | 1380 |
| ٠ | TCTTTGTACC | AAGTCAACAA | TCATATACCA | GTAGTCAAAT | TGAAAATATG | AATGCATGGT | 1440 |
| | TAAAACCTTA | TCGACGCGCA | TTACATTTAG | AAGAGCCGGT | TGTAAAAATT | ACCGAAGAAC | 1500 |
| 10 | AACAATTTAA | GTGTCATATT | GATACGATAA | TTAAAAAGCA | AGTGCCTGTA | TGTTGTTTTA | 1560 |
| | CTTTTGGAAT | TCCAAGCGAA | CAGATTATAA | GCAGGTTGAA | AGCAGCGAAT | GTCAAACTTA | 1620 |
| 15 | TAGGTACAGC | AACAAGTGTT | GATGAAGCTA | TTGCGAATGA | AAAAGCGGGT | ATGGATGCTA | 1680 |
| | TCGTTGCTCA | aggtagtgaa | GCAGGTGGAC | ATCGTGGTTC | ATTTTTAAAA | CCTAAAAATC | 1740 |
| | AATTACCTAT | GGTTGGAACA | ATATCTTTAG | TGCCACAAAT | TGTAGATGTC | GTTTCAATTC | 1800 |
| 20 | CGGTCATTGC | CGCTGGTGGA | ATTATGGATG | GTAGAGGAGT | TTTGGCAAGT | ATTGTCTTAG | 1860 |
| | GTGCAGAAGG | GGTACAAATG | GGCACCGCAT | TTTTAACATC | ACAAGACAGT | AATGCATCAG | 1920 |
| | AACTACTGCG | AGATGCAATT | ATAAATAGTA | AAGAAACAGA | TACAGTCATT | ACAAAAGCGT | 1980 |
| 25 | TTAGTGGAAA | GCTTGCACGC | GGTATCAACA | ATAGGTTTAT | CGAAGAAATG | TCCCAATACG | 2040 |
| | AAGGCGATAT | CCCAGATTAT | CCAATACAAA | ATGAGCTAAC | AAGTAGCATA | AGAAAAGCCG | 2100 |
| 30 | CAGCAAACAT | CGGCGACAAA | GAGTTAATAC | ATATGTGGAG | TGGACAAAGC | CCGCGACTAG | 2160 |
| | CAACAACGCA | TCCCGCCAAC | ACCATCATGT | CCAATATAAT | CAATCAAATT | AATCAAATCA | 2220 |
| | TGCAATATAA | ATAATCGACC | GCAATCCACA | AAAGCACAAG | CACCCCCAAA | CATTATTTTA | 2280 |
| 35 | GTGCTTGCCA | TTTTTGTGGA | TTGCGTTTCT | ATTTTACCAA | TTTAATCAAA | CGAAAACATC | 2340 |
| | AAGCTGAAGA | TCGCCGAAAG | ATTTTAATCA | AGCAAAAACA | TCAAACTAAA | GTTCGCTGAA | 2400 |
| | ATGATTATGA | TAAAAGTTAT | ATGGTATGAT | GACATTGGTG | ATATATATGA | TAAACATCGG | 2460 |
| 40 | ATTAACAGGT | TGGGGTGATC | ACTATTCATT | ATATGAAGAT | TTAGAACGCC | AAACCGATAA | 2520 |
| | ACTTAAAACA | TATGCTGGAC | ATTTTCCGGT | TGTCGAATTA | GATGCGACAT | ACTATGCGAT | 2580 |
| 45 | ACAACCGGAA | AGAAATATAT | TGAAATGGAT | AAAAGAAACG | CCTGATACAT | TTGAATTTGT | 2640 |
| 45 | GGTCAAAATT | CATCAAGCAC | TCACATTGCA | TGCAGACTAC | AAAACATTTG | CAGATACAAG | 2700 |
| | GCAAGAACTA | TTTGATCAAT | TTAAGAATAT | GTTAGAGCCC | TTACATACAC | AGAAAAATT | 276 |
| 50 | AGCAATGGTA | TTGGTTCAAT | TTCCGCCATG | GTTTGACTGC | AATGCACAAA | ATATCAAATA | 282 |
| | TATTTTGTAT | GTAAGACAGC | AATTACAAGC | ATTTCCAATG | TGTGTAGAAT | TTAGGCATCA | 288 |
| | ATCATGGTTT | AGTGATGCAT | TTAAAGAACA | AACATTGGCA | TTTTTAACAG | AACATCAAAT | 294 |

| | AATCACAAAT | GAAATTGCGT | TTGTACGTTA | TCATGGACG | r aatcattacg | GTTGGACTAA | 3060 |
|----|------------|------------|--------------|------------|--------------|------------|------|
| | GAAAGATATG | TCAGATCAAG | AATGGCGCGA | TGTACGCTA: | TTATATGATT | ATAATGAGCA | 3120 |
| 5 | AGAATTAATA | GACTTGGCAC | : AAAAGGCACA | AATATTAGC | A CAAAAAGCTA | AGAAAGTTTA | 3180 |
| | CGTCATATTT | AACAATAATT | CTGGTGGTCA | TGCAGCAAA1 | RATGCCAAAA | CATATCAGCG | 3240 |
| 10 | ATTATTGAAT | ATAGAATATG | AAGGGTTAGC | ACCACAACAA | TTAAAATTAT | TTTAAGAGGC | 3300 |
| | GACGACTATG | TTATTAACAA | TTACATTATT | AGTTTTAATO | GGAGGTTTGT | CAGCGATTAT | 3360 |
| | AGGGTCTATC | GTAGGCATTG | GAGGCGGTAT | TATTATCGTT | CCAACAATGG | TTTACCTCGG | 3420 |
| 15 | TGTTGAACAT | GGATTACTAC | ATAATATTAC | AACACAAGTA | GCGATAGGGA | CGTCTTCAGT | 3480 |
| | CATTCTAATT | GTGACAGGAC | TTTCTTCATC | ACTTGGATAT | TTAAAAACAA | AACAAGTTGA | 3540 |
| | TATTAAAAAT | GGTTCCATCT | TTTTATTTGG | ACTATTACCA | GGTTCATTGC | TTGGGTCCTT | 3600 |
| 20 | CATTAGTAGA | TATTTAACAT | TTGAGTCATT | TAATTTATAT | TTTGGTATCT | TTTTAATTTT | 3660 |
| | CGTAGCCATT | TTATTAATGG | TAAGAAATAA | GATTAAACCG | TTTAAAATTT | TCGATAAACC | 3720 |
| 25 | CAAGTATGAA | AAGACTTATG | TAGACGCTAA | AGGTAAAACA | TATCATTATA | gTGTTCCACC | 3780 |
| :5 | ATTGTTTGCT | TTTATTACAA | CGTTTTTAAT | TGGTATATTG | ACAGGTTTAT | TTGGTATTGG | 3840 |
| | AGGTGGCGCA | CTAATGACGC | CACTAATGCT | TATTGTATTT | AGATTTCCAC | CTCATGTAGC | 3900 |
| 10 | TGTTGGAACA | AGTATGATGA | TGATTTTCTT | TTCAAGTGTC | ATGAGTTCTA | TAGGGCACAT | 3960 |
| | TGCTCAAGGT | CACGTAGCTT | GGGGTTATGC | AATCATnTTA | ATTATTTCTA | GTTATTTTGG | 4020 |
| | TGCGAAAATC | GGTGTCAAAG | TGAATCAATC | AATTAAGTCA | GATACGGTAG | TAACATTATT | 4080 |
| 15 | GAGAACAGTA | ATGTTGTTAA | TGGGTATATA | TTTAATTATT | CGTGCGTTGA | TTTAATACAA | 4140 |
| | CTTTAAAAGG | AGGACGTCAA | TTTGAGGCTT | ACAATTTATC | ATACGAACGA | TATTCATAGT | 4200 |
| | CATTTACATG | AATACGAACG | CATTAAAGCA | TATATGGCAG | AACATCGGCC | ACGACTTAAT | 4260 |
| | CATCCTTCTT | TATATGTTGA | TCTAGGTGAT | CATGTAGATT | TATCCGCACC | TATAACTGAA | 4320 |
| | GCAACTTTAG | GTAAAAAGAA | TGTGGCATTA | CTAAATGAAG | CAAAATGTGA | TGTTGCAACA | 4380 |
| 5 | ATCGGTAATA | ATGAAGGGAT | GACCATTTCA | TACGAAGCTT | TAAATCACCT | TTACGACGAA | 4440 |
| | GCAAAATTTA | TAGTGACATG | TAGCAATGTT | ATAGATGAAT | CAGGTCATTT | ACCAAATAAT | 4500 |
| | ATCGTTTCTT | CTTATATTAA | GGACATAGAC | GGTGTGAAAA | TACTATTCGT | TGCAGCGACA | 4560 |
| 0 | GCACCTTTTA | CCCCATTTTA | TCGTGCACTA | AATTGGATTG | TTACCGATCC | ACTTGAATCT | 4620 |
| | ATAAAAGAAG | AAATTGAACT | TCAACGAGGT | AAATTTGATG | TATTAATCGT | GCTAAGTCAT | 4680 |
| | TGTGGCATTT | TCTTCGATGA | AACATTATGC | CAAGAATTGC | CTGAAATTGA | TGTCATTTTT | 4740 |

| | GCAGCTGGAA | AGTATGGTAA | TTATCTTGGA | GAGGTTAATT | TAACTTTTGA | GGCACATAAA | 4860 |
|----|---------------------|------------|------------|------------|------------|------------|------|
| | GTAGTACATA | AAACTGCAAA | GATTATTCCT | TTAGAAACAT | TACCTGAAGT | TGAAACTTCA | 4920 |
| 5 | TTTGAAGAAG | AAGGAAAAAC | GTTAATGTCC | AATTCAGTAA | TTCAACATCC | AGTAGTGCTT | 4980 |
| | AAGCGTAGTA | TGAATCACAT | AACTGAAGCT | GCATACTTAT | TAGCTCAAAG | TGTTTGTGAG | 5040 |
| | TATACACATG | CACAATGTGC | CATCATCAAT | GCTGGCTTAC | TCGTTAAAGA | TATTGTAAAA | 5100 |
| 10 | GATGAAGTGA | CAGAATATGA | CATTCATCAA | ATGTTACCGC | ATCCGATTAA | TATGGTAAGG | 5160 |
| | GTTAGACTTT | TTGGTGTGAA | ATTAAAAGAG | ATTATAGCTA | AAAGTAATAA | ACAAGAATAT | 5220 |
| 15 | ATGTATGAAC | ATGCACAAGG | TTTGGGTTTC | AGAGGGAATA | TATTTGGAGG | ATATATTCTT | 5280 |
| | TATAATTTAG | GGTACATTCA | TTCTACAGGG | CGTTACTATC | TGAATGGAGA | AGAAATCGAA | 5340 |
| | GACGACAAAG | AATATGTACT | AGGTACGATA | GATATGTATA | CGTTCGGTCG | TTATTTCCCA | 5400 |
| 20 | ACATTGAAAG | AATTACCAAA | AGAGTATTTA | ATGCCAGAGT | TTTTAAGAGA | TATATTTAAA | 5460 |
| | GAAAAATTAT | TGGAATATTA | AAAAGTAAGA | TTATTGGATT | TTCATTTGTC | ATGAATTTCG | 5520 |
| | ATATAATGTT | TAAAGATACA | CTTAACAGGA | GGGTATGTGT | TGTTATGGCG | ACAAAAAACG | 5580 |
| 25 | AGGAAATATT | ACGTAAACCG | GATTGGTTGA | AAATAAAATT | AAATACCAAC | GAAAACTATA | 5640 |
| | CAGGACTTAA | GAAGATGATG | AGGGAAAAA | ATCTTAATAC | TGTATGTGAA | GAAGCTAAAT | 5700 |
| 30 | GTCCTAATAT | ACATGAATGT | TGGGGTGCAC | GTCGTACAGC | GACATTTATG | ATTTTAGGTG | 5760 |
| | CCGTATGTAC | AAGAGCTTGT | CGTTTTTGTG | CGGTTAAGAC | AGGTTTACCT | AATGAACTTG | 5820 |
| | ATTTAAATGA | GCCTGAACGT | GTAGCTGAAT | CAGTTGAATT | AATGAATTTG | AAACACGTTG | 5880 |
| 35 | TTATCACTGC | TGTTGCGCGT | GATGATTTAA | GAGATGCTGG | TTCAAATGTT | TATGCTGAGA | 5940 |
| | CAGTACGTAA | AGTTAGAGAA | AGAAATCCAT | TTACAACGAT | TGAAATTTTA | CCATCAGATA | 6000 |
| | TGGGCGGGA | CTATGATGCG | TTAGAAACAT | TAATGGCGTC | AAGACCTGAC | ATTTTAAACC | 6060 |
| 40 | ATAATATTGA | AACTGTTCGT | CGCTTAACAC | CGAGAGTTCG | TGCGCGTGCG | ACTTACGACA | 6120 |
| | GAACATTAGA | GTTTTTACGT | CGTTCAAAAG | AATTACAACC | GGATATCCCA | ACTAAATCAA | 6180 |
| | GTATTATGGT | TGGATTAGGT | GAAACTATAG | AAGAAATTTA | TGAAACGATG | GATGATTTAC | 6240 |
| 45 | GTGCGAATGA | TGTAGATATT | TTAACGATTG | GTCAATATTT | ACAACCTTCA | CGTAAACATT | 6300 |
| | TAAAGGTTCA | AAAATATTAC | ACGCCTTTAG | AGTTTGGTAA | ATTAAGAAAA | GTGGCAATGG | 6360 |
| 50 | ATAAAGGGTT | TAAACATTGC | CAAGCTGGAC | CTTTAGTACG | TAGTTCTTAT | CATGCGGATG | 642 |
| | AGCAAGTAAA | TGAAGCTGCT | AAAGAAAAGC | AACGCCAAGG | TGAGGCACAC | TTAAATAGTT | 648 |
| | እ አ ጥአጥጥ እአር | | CCATAAACCC | ттьсттсть | CAAAACGAAC | GTGTCATAGA | 654 |

| | AGGTGAAGAA | TTTGATAAAA | GTAGATCAAC | ATTACTTTGA | ATTAATAGAA | AATTATCGCG | 6660 |
|-----|------------|------------|------------|------------|------------|------------|--------------|
| | AATGTTTTAA | TGAAGAACAA | TTTATTGCTA | GGTATTCAGA | TATTTTAGAT | AAATATGATT | 6720 |
| 5 | ACATAGTTGG | TGACTATGGT | TACGATCAAT | TACGATTAAA | AGGTTTTTAC | AAAGATTCTA | 6780 |
| | ATAAAAAAGC | AGAGATGAGT | AAACGTTTTT | CAAATATTCA | AGATTACATA | TTTGAATATT | 6840 |
| 10 | GTAACTTTGG | TTGTCCTTAC | TTTGTATTAA | GACATTTGTC | TAAACAAGAG | GTTAAAAAGT | 6900 |
| ,,, | TAATCGAAGA | AGTTCATCCG | TCTGATGTGA | TAGATGACGA | CAATAAACTT | CAAGATGTGA | 6960 |
| | AGATTAAGCC | AACCATTCAA | GATACTGAAC | ATTAATAAAA | CCCTTAGCTA | GATTGAAAAT | 7020 |
| 15 | GGGAATCATG | CAATTCAAGC | ATGGACCTGT | AATCTAGTTA | GGGGTTTTTA | TCTTTAATGA | 7080 |
| | ATGACTTCAT | TTAAATACTC | AGTAATTTCA | TCGCCTTCTT | CAGCATTTAC | ACCTAAAATA | 7140 |
| | TGAGCGATAT | AGCCTTCTTC | TTTTAAATCA | TCAGTACCGA | TAATACCGAA | TTTATTTGTT | 7200 |
| 20 | TGCATATTAA | GTACGAGTGT | CTTACCATAA | TGTCTATTTG | TATGGACTAA | CATCAAATCA | 7260 |
| | TATCGACTAT | GCTCGCCAAC | AAAACCAACA | AACTGAACTT | GACTCTCTTC | GTTGTCATCA | 7320 |
| | TATAAATACA | TATCAATCAT | TTTGTAGCGA | CTCCTTTTAA | AAGTAGTAAA | GTTAGTATAA | 7380 |
| 25 | CGACAAATGA | AGTATACTGC | AAAATTATGA | TAATATATAA | gtgagaggtg | ACAAGGAATG | 7440 |
| | TATTTTGTAG | ACAAAGATAA | ACTAACTCAG | AAATTAGCCT | ATTTACAAGC | ATTAACTGAT | 7500 |
| 30 | GATTATCATG | AGAGCAAGCA | CAATCATTAT | GCATTTGAAC | GCATTGCTCA | AATGTTGATA | 7560 |
| | GAATCATCGG | TAGATATAGG | GAATATGATT | ATCGATGCAT | TTATTTTAAG | GGATCCTGGT | 7620 |
| | AATTATAAAG | ATGTGATTGA | TATATTAGAA | CTAGAAAATG | TTATTACTAA | AGAAACACAG | 7680 |
| 35 | CAGGCGATTA | ATAAAACTGT | CGGTATTCGT | AAACAATTTA | CATATGATTA | CACAGCCTTA | 7740 |
| | GATGTTGAGA | TTATCATGCC | AATGTTTGA | | | | 7 769 |

(2) INFORMATION FOR SEQ ID NO: 215:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 644 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(D) TOPOLOGY: Timea.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 215:

ACCGCCACCC ATTAATGATT GCTTAAAATC AATAGTCGTA CCATTTAATA CGGGTGCATC 60
TTTTTTGTCT ACTAATACTT TTAATCCAAA GTATTCTAAG ACTTCATCAT TTTCACCAGG 120
CGCTTCTTCT GCACCCATAC CGTATGTTAA ACCAGTGCAC CCGCCACCAT TCACTTTAAT 180

55

| | | |
|----|--|-----|
| | TGCTTCTGTT AATATAACTG TTGGCATGAT AACTCCTCCT TAAAAAAATCC AAGTTTCTTT | 300 |
| | TATATGTGCA TATATATTT GTAATAATTC TTCCGGCGAA TCACCTTCAA CAATATCACC | 360 |
| 5 | ATTTACTAAA GCATACAACC CGGCTGAACA TATACCACAA TGTGTCAGGC AACCATACTC | 420 |
| | TAACACATCG ACATCTGGGT CATTTTCCAG TTGATTAAAA ACATAATCTC CACCTTTTGC | 480 |
| 10 | CATGTTAGAG AGACAAAATT CTACGATCGG ATTCATACTT CACCTTCTTA TTTCATTTGT | 540 |
| | TACAATATTA TAGCATTTTA AAACTGGTAT TTTAACATGA TGTGCTCAAT TAGCAACAAC | 600 |
| | TGATGTTTCT TATCCCAGTT ATGTAATAGT GCCTTAGTTA GTAC | 644 |
| 15 | (2) INFORMATION FOR SEQ ID NO: 216: | |
| | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1578 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double | |
| 20 | (D) TOPOLOGY: linear | |
| | | - |
| 25 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 216: | . • |
| | GAATGATGAA AGGAATAGAT AAATAATGTA ATAGATTTAT CCGAGAAAAT | 60 |
| | TGAAAGAACA AAAGATATGC CAATCAAGAA TACTATAACT ACTCAATTAG GAAATAAACT | 120 |

GAATGATGAA AGGAATAGAA AAGAAAAGAT AAATAATGTA ATAGATTTAT CCGAGAAAAT
TGAAAGAACA AAAGATATGC CAATCAAGAA TACTATAACT ACTCAATTAG GAAATAAACT
TATTGGCACA AAAAAAGCTC GTTTTGATGA TAAGAAAGTA GTGTCGTTTG GAGCATTTGA
AGATGAATAA AATAAATGAT AGAGATTTAA CAGAATTGAG TAGCTATAGG GTTTATCAAG
ACATCAATAA AGATAATGAC TITACAGTTA ACGAAAAACG ATTTAAGCAG GCAGATGTAT
TTGAAGATTT ATATAGAGAG AAACTAAAAG ACAAAATAA ATTAAGAAGG TATAATTATT
TACAAAATGA AACTTTTAAA AGCGCATAAA TAGGTGATGA GATATGCTTA AAAAAGCAAA
ATTTATCTTA ATGGCAACGA TACTACTATC AGGATGTTCA ACTACCAATA ACGAATCCAA
CAAAGAAACA AAATCTGTAC CAGAAGAAAT GGATGCTTCA AAATATGTAG GACAAGGATT
CCAACCACCT GCAGAAAAAAG ATGCGATTGA ATTTGCAAAG AAGCATAAAG ATAAAATTGC
TAAGGGAGGC GAACAATTT TTATGGATAA CTTCGGTCTA AAAGTTAAAG CTACAAATGT
TATAGGTAGT GGCGATGGTG TAGAAGTATT CGTGCATTGT GATGACCACG AYATCGTATT
TAATGCGAGT ATTCCATTTG ATAAATCAAT WATTGASAGT GATAGCTCAT TAAGAAGTTA
GGAYAAAGGY GATGATATGA GTACTTTAGT TGGTGCAGTA CTCAGTGGGT TTGAATATCG

AGCACAAAAA GAAAATATG ATAAATTATA TAAATTTTTC AAAGATAATG AAGAGAAATA

TCAATATACA GGATTTACAA AAGAAGCAAT TAATAAGACG CAAAATAGTG GTTATGAAAA

| ACCATTGTTA | AACAAAAGTG | ACAGTGAATT | TTCAAAAGAA | TTGTCAAATG | TTAAGAAGCA | 1080 |
|------------|------------|------------|------------|------------|------------|------|
| ATTAAAAGAT | AAGTCTAAAG | TTTCGGTAAC | TACTACTCTA | TTTAGTAAAA | AAAAGAACTA | 1140 |
| TACTAAAAA | AGTAACAGTG | AAAATGTAAT | AAAAATGGCA | GAAGAAATAA | AAAAGATAA | 1200 |
| AGAGATACCA | AACGGTATAG | AGCTTAGTAT | AAAATTTTCG | GACAATAAAA | TAAATACGGT | 1260 |
| TAAACCAAAT | TTTAACGGTG | aAAGCACTTC | AGAATATGGT | GTGTTTGATC | AAGAATAAAA | 1320 |
| TTAATGATGa | AAATTTAACG | GAGAATAGTG | TATATTGAGT | AGATCMAGAA | TAAAAAGATA | 1380 |
| ATTCTACTAT | TGTTGTGAAG | GCAAATAAGT | AGAAGATTTT | AAGTGTAATT | TCTGGTGATT | 1440 |
| AATAATAA | TATAnATGGn | AGTACTGATA | TAAnACTTTT | TAACCTACTA | GATTCTTATA | 1500 |
| ATTTGCTTTC | CATTTTATGA | CGATTTTTAC | TCCAATTGAG | TGATAGAATC | CAAAAAAGCC | 1560 |
| ATCTCCAAAA | ATTAATCC | | | | | 1578 |

(2) INFORMATION FOR SEQ ID NO: 217:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5137 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 217:

TGTTTTCCTT GGGTTAAAAC ATGCTTGCTA TGCGTTTGTA AATATGACTT GCTGTTTTnA 60 CCTGnATACC CGTCACACCA TGGAAGTAAA AATGTTTCTT GCTCTTGGCT TACAATTTTA 120 GCTTTAATCG CTTCATATGC TTTATATTGG TCTTCTGTTA ATTGCTGTTT TGATTCTTGT 180 TCGAAAACAC GATCTTTAAA TGGGTCTCTT TCAACAACCG CGTCATATTT TTCAACATAA 240 CCTTFTTTGA TAAGTCCATC TAAACTGGAT TTTGAAAAGC CCATATCCTC AATATCAGTT 300 AAAAATATTG TTTTATGTTG TTCTTCAGAC AAGTAAGCAT ACAAATCGTA TTGTTTAATA 360 ACTITCICCA ACTIAGCIAA TACTICATCA GGATGATACC CITCAATGAC ACGAACAGCA 420 CGCTTGGTTT TTTTAGTTAT ATTTTGTGTG AGAATCGTTT TTTCTTCAAC GATATCATCT 480 TTTAACAACT TCATAAGCAA TTGAATATCA TTATTTTTTT GCGCATCTTT ATAATAATAG 540 TAACCATGCT TATCAAATTT TTGTAATAAA GCTGAAGGTA GCTCTATGTC ATCTTTCATC 600 TTAAATGCTT TTTTATACTT CGCTTTAATA GCACTCGGAA GCATCACTTC TAGCATAGAA 660 ATACGTTTAA TGACATGAGT TGAACCCATC CACTCACTTA AAGCTATTAA TTCTGATGTT 720 AATTCTGGTT GTATATCTTT CACTTCTATG ATTTTTTTTA ACTTCGAAAC GTCAAGTTGT 780

5

10

15

20

25

35

40

45

50

| 5 | ACAATTACAC | GCACACCAGG | TTGGATGACA | GATTCGAGTT | GTTCGGGAAT | TATATAATCA | 900 |
|-----------|------------|------------|------------|------------|------------|------------|------|
| | AATTTATAGT | CAACGCTCTT | CGACGCGACA | TCGACTATGA | CTTTCGCTAT | CATTATTGCC | 960 |
| | ACCTAGTTTC | TAGTTCATCT | AAAATTTGTG | CAGCTAATAC | TACTTTTTT | CCTTTCTTGA | 1020 |
| | TATTTACTTT | TTCATTATTT | TTAAAATGCA | TTGTCAATTC | ATTATCATCA | GAACTAAATC | 1080 |
| | CGATAGACAT | ATCCCCAACA | TTATTTGAAA | TAATCACATC | TGCATTTTTC | TTGCGTAATT | 1140 |
| | TTTGTTGTGC | ATAATTTTCA | ATATCTTCAG | TCTCTGCTGC | AAAGCCTATT | AAATACTGTG | 1200 |
| | ATGTTTTATG | TTCACCTAAA | TATTTAAGAA | TGTCTTTAGT | ACGTTTAAAA | GATACTGACA | 1260 |
| 15 | AATCACCATC | CTGCTTTTTC | ATCTTATGTT | CTAATACATC | AACCGGTGTA | TAGTCAGATA | 1320 |
| | CGGCTGCTGC | TTTTACAACA | ATATCTTGTT | CGTCAAATCG | GCTTGTCACT | TGTTCAAACA | 1380 |
| | TTTCTTCAGC | ACTTTGAACA | TGAATAACTT | CAATATCTTT | TGGATCCTCT | AGTGTTGTAG | 1440 |
| 20 | GACCAGCAAC | TAACGTCACG | ATAGCTCCTC | GATTTCGCAA | TGCTTCAGCT | ATTGCATAGC | 1500 |
| | CCATTTTTCC | AGAAGAACGA | TTGGATACAA | ATCTGACTGG | ATCGATAACT | TCAATAGTTG | 1560 |
| 25 | GTCCTGCTGT | AACCAATGCG | CGTTTATCTT | GAAATGAACT | ATTAGCTAAA | CGATTACTAT | 1620 |
| | TTTGAAAATG | AGCATCAATT | ACAGAAACGA | TTTGAAGCGG | TTCTTCCATA | CGTCCTTTAG | 1680 |
| | CAACATAACC | ACATGCTAGA | AATCCGCTTC | CTGGTTCGAT | AAAATGATAC | CCATCTTCTT | 1740 |
| 30 | TTAAAATATT | AATATTTTGC | TGCGTACGTT | TATTTTCATA | CATATGCACA | TTCATAGCAG | 1800 |
| | GCGCAATAAA | TTTCGGTGTC | TCTGTTGCTA | GCAACGTTGA | TGTCACCAAA | TCATCAGCAA | 1860 |
| | TACCTACACT | CAATTTTGCA | ATTGTATTTG | CCGTTGCAGG | TGCAACAATG | ATTGCATCTG | 1920 |
| 35 | CCCAATCACC | TAATGCAATA | TGCTGTATTT | CTGAAGGATT | TTCTTCTATA | AAAGTATCTG | 1980 |
| | TATAAACAGC | ATTTCGACTT | ATTGCTTGAA | ATGCTAATGG | TGTCACAAAT | TTTTGTGCGT | 2040 |
| 10 | GATTÉGTTAA | CATAACGCGA | ACTTCATACC | CAGATTGTGT | TAACTTACTT | GTCAAATCAA | 2100 |
| 40 | TTGCTTTATA | TGCCGCAATG | CCACCTGTAA | CGGCTAATAA | TATTTTCTTC | ATATTCAATC | 2160 |
| | TCCCTTAAAT | ATCACTATGA | CATTTACGCT | TTACATCATC | ATATGCGCAC | AAATGCTCAT | 2220 |
| 45 | TACTTTTTTA | TAGATACAAA | TTTAGTATTA | TTATAACATC | AATCATTGGA | ТАААСТАААА | 2280 |
| | AAACACACCT | ACATAGGTGC | GTTTGATTTG | GATATGCCTT | GACGTATTTG | ATGTACGTCT | 2340 |
| 50 | AGCTTCACAT | ATTTTTAATG | GTCGAAACTA | TTCTTTACCA | TAATAATCAC | TTGAAATAAC | 2400 |
| | AGGGCGAATT | TTACCGTCAG | CAATTTCTTC | TAACGCTCTA | CCAACTGGTT | TAAATGAATG | 2460 |
| | ATATTCACTT | AATAATTCAG | TTTCAGGTTG | TTCATCAATT | TCACGCGCTC | TTTTCGCTGC | 2520 |
| | AGTTGTTGCA | ATTAAATACT | TTGATTTAAT | TTGTGaCGTT | aATTGGTTtA | AAgGTGGATT | 2580 |
| <i>55</i> | | | | | • | | |

| | TTTATGTGC | r CAGCTTCTAC | AATACATTG | A ATTCEATTCY | TCGcAAGtT | TACTTCAtCA | 270 |
|----|------------|--------------|------------|--------------|------------|------------|-------|
| 5 | TTAACTACA | A CGTAAYCGTA | TAAATTCATC | ATTTCTACTI | CTRTACGCGC | YTCGTTAATA | 276 |
| | CGACTTTGT | A TTTTCTCATC | AGATTCTGTT | CCTCTACCTA | CTAATCGCTC | TCTCAAGTGT | 2820 |
| | TCTAAACTTC | GAGGTGCTAA | GAAAATAAAT | AGCGCATCTG | GAAATTTCTT | TCTAACTTGC | 288(|
| 10 | TTTGCACCTT | CTACTTCAAT | TTCTAAAAA1 | ACATCATGAC | CTtCGTCCAT | TGTATCTTTA | 294(|
| | ACATATTGA | CTGGTGTACC | ATAATAGTTG | CCTACATATT | CAGCATATTC | TATAAATTGG | 3000 |
| | TCATCTTTGA | TTAAAGCTTC | AAACGCATCC | CTAGTTTTAA | AAAAGTAATC | TACGCCATCA | 3060 |
| 15 | ACTTCACCTT | CACGCATTTG | ACGTGTTGTC | ATTGAAATAG | AATACTTATA | TGATGTACTT | 3120 |
| | GGATCTTCAA | ATATRCGTRT | TCTAACAGTA | CCTTTACCTA | CTCCAGATGG | TCCTGATAAA | 3180 |
| | ACGATTAACA | ATCCTTTTTC | ATTATCCATG | CCTTACGACC | TCTCTAAGCT | AATCTTCTAT | 3240 |
| 20 | TATTTAAATA | TGATATCACA | TTGTTCTTTA | TATTGTATAG | CATATTTGAA | ATTGCATGCC | 330.0 |
| | ATAATTTCTA | TTAAGTCTAA | CAATATCGTT | ATATTGCACG | ATTAATTTTA | ATTAAATAAA | 3360 |
| 25 | TTGAATTGCA | AACTTTTAGA | TAATGTAAAA | TGTATGGCAT | AATGTATGGT | TCAATAACTA | 3420 |
| | TACTGAAAAG | TTACAATCAT | GTTAAAATGA | AACGAATGAT | ATGAAGAAGG | TGGAAGATAA | 3480 |
| | ATTATGGCTT | ATGATGGCTT | ATTTACAAAG | AAAATGGTTG | AGTCTCTACA | ATTTTTAACA | 3540 |
| 30 | ACAGGACGTG | TTCACAAAAT | CAATCAACCT | GATAATGACA | CGATACTAAT | GGTTGTACGT | 3600 |
| | CAAAATAGAC | AAAACCATCA | ATTGTTATTG | TCAATCCATC | CAAACTTTTC | AAGATTACAA | 3660 |
| | TTGACTACTA | AAAAATATGA | TAATCCATTT | AATCCACCCA | TGTTTGCGCG | TGTTTTTAGA | 3720 |
| 35 | AAACACTTAG | AAGGTGGTAT | TATCGAATCG | ATTAAGCAAA | TTGGTAATGA | TCGTCGCATT | 3780 |
| | GAAATCGATA | TAAAGAGTAA | AGATGAAATT | GGCGATACTA | TTTACCGCAC | TGTCATCCTT | 3840 |
| 40 | GAGATTATGG | GTAAACATAG | TAACTTAATT | TTAGTAGATG | AAAATCGCAA | AATAATTGAA | 3900 |
| | GGATTTAAAC | ACTTAACACC | AAATACGAAT | CACTATCGTA | CAGTAATGCC | AGGATTTAAT | 3960 |
| | TATGAAGCAC | CACCTACTCA | GCACAAAATA | AATCCGTATG | ATATTACAGG | TGCAGAGGTG | 4020 |
| 45 | TTGAAATATA | TCGATTTTAA | CGCAGGTAAT | ATTGCTAAAC | aattattgaa | TCAGTTTGAA | 4080 |
| | GGATTTAGCC | CTTTAATTAC | GAATGAAATC | GTTAGTCGTC | GTCAATTTAT | GACTTCATCA | 4140 |
| | ACATTACCAG | AAGCATTTGA | CGAAGTAATG | GCAGAAACCA | AGTTACCACC | TACTCCTATT | 4200 |
| 50 | TTTCATAAAA | ATCATGAAAC | AGGTAAAGAG | GATTTCTATT | TTATAAAGTT | AAATCAATTT | 4260 |
| | AATGATGATA | CAGTTACATA | CGATTCATTA | AATGATTTGC | TTGATCGTTT | TTATGATGCG | 4320 |
| | CGTGGCGAAC | GTGAACGCGT | TAAACAACGT | GCGAATGATT | TAGTTCGATT | TGTTCAACAG | 4380 |

| ATAAAGATAC | TGAACAGTTA | TATGGTGAAT | TGATCACTGC | TAATATAT | CGAATTAAGC | 4500 |
|------------|------------|------------|------------|------------|------------|-------|
| AAGGCGATAA | AGAAGTGACG | GCATTGAATT | ATTATACGAA | TGAAGAAGTT | GTCATTCCTT | 4560 |
| TAAATCCTAC | AAAATCCCCA | TCAGCAAATG | CTCAATATTA | TTATAAACAA | TATAAyCGTA | 4620 |
| TGAAAACGAG | AGAmCGTGAA | TTACAACATC | AAATTCAATT | GACGAAAGAC | AATATAGATT | 4680 |
| ATTTTTCAAC | AATCGAACAA | CAATTACATC | ATATTTCTGT | CCATGACATT | GATGAAATTA | 4740 |
| GAGATGAATT | AGCAGAACAA | GGCTTTATGA | AACAGCGTAA | AAATCAAACT | AAGAAAAGA | 4800 |
| AAGCGCAGAT | TCAATTACAA | CATTATGTAT | CAACTGATGG | CGACGATATA | TATGTTGGTA | 4860 |
| AGAATAACAA | GCAAAATGAT | TATTTAACAA | ATAAAAAAGC | TAAAAAAACT | CACACATGGT | 4920 |
| tACACACAAA | AGATATTCCT | GGTTCACATG | TCGTTATATT | TAATGATGCA | CCAAGTGATA | 4980 |
| CGACAATCAA | GGAAGCGGCT | ATGTTAGCAG | GATACTTTTC | AAAAGCTGGT | AATTCTGGAC | 5040 |
| AAATACCTGT | TGATTATACA | ттааттаааа | ATGTGCATAA | ACCATCaGGT | GCAAAGCCTG | 5100 |
| GGTTTGTAAC | ATATGACAAT | CAAAAAACTT | TGTATGC | | | 513,7 |
| | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 218:

5

15

20

25

30

35

40

50

55

- . (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2267 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 218:

| GTTTTATCGC | AGCAGTAAAG | CTATCAATCG | GCGGTTCAAT | TGATGATGCA | TTAGCAGAAA | 60 |
|--------------------|------------|------------|------------|-------------------|------------|-----|
| TCAnacaato | ATTTTAGTTA | AAATTTACTA | ATAATGAAAA | ATGTAAACCT | TTTTCAAATG | 120 |
| АААСТТТАТ а | AAAAATATGA | TAGTATATAT | GTAAATGTTT | AATAAAATCT | GGAGAAATAG | 180 |
| GAGGACATTG | CCATGCAACA | CCTTATAAAA | AAACATGTAT | TGAATGGCGA | GTTTGATTTA | 240 |
| GTACGACAAT | TGATGTCCGA | AACAGATTTT | ATGGAATTTG | AAGAAGCATA | TATTTCAAGT | 300 |
| GCGCATGAAG | TAGAAAGTAT | GATGTTTTAT | ACATGTATTT | TAGATATGAT | TAAGTACGAA | 360 |
| GAATCATCTG | AAATGCATGA | CTTAGCATTT | TTATTGCTTG | TGTATCCACT | AAGTGAATAT | 420 |
| GAAGGTGCTI | TGGATTCTGC | TTATTATCAT | GCAGACGCTT | CCATAAAACT | TACTGACGGC | 480 |
| AAAGAAGTTA | AAAGTTTGTT | ACAAATGTTA | TTATTGCATG | CGATACCAAC | ACCTGTTATT | 540 |
| TCAGATAAGA | AGGCTTTTGA | TATCGCCAAG | CAAATTTTAA | AATTAGATCC | TAATAATAAT | 600 |
| GTTGCTCGTA | ACGTCTTAAA | AGACACTGCC | AAACGTATGC | GACAACGTTG | TTGTTGATAT | 660 |

| | AGIIITAACA TITGGTTGGG TTGGGCATAT GTTCCAGCCT TTTTTAATAC TTAAAAACTA | 780 |
|-----------|---|------|
| 5 | ACGAAGTATA CTTGTGTGCA CAAATGGTTT TTATACAACA TTTTATAAAT TTATACATTT | 840 |
| | TAATAAAGAA CATACGATAG ATGGTTTAAA CCTTGTTAAC TGAGAAATTT TGATATGTAT | 900 |
| | TCTTCGAAAT TTAACTAAAT ATACGAAATT CAAGAAGCAC AATAATTAAT CATTTTTCCT | 960 |
| 10 | ATACAAAAGT TCGTATGACT GCATTATAAA AGCATAAATT TATAATTTTT TTAAATGTCA | 1020 |
| | TTGAACGTGA TAATGTGAAT GGATTGAGCA ATTTTGAAAA AGTGAAAAAT AACCTATGCG | 1080 |
| | ACTTGCAATT AATTTTCAGT ACGTTATAAT GCACACTGTG CAAAATTAAG GAGGTCTATT | 1140 |
| 15 | ATTCACATGA TGATGAATAA AGAAGCAACA AAAATTGGAT TTGCCTACGT CGGCATTGTA | 1200 |
| | GTGGGCGCAG GATTTTCAAC TGGACAAGAA GTTATGCAAT TTTTCACTAA ATATGGCTTG | 1260 |
| 20 | TGGGCTTATT TAGGTGTTAT TATATCTGGT TTTATTTTAG CTTTTATTGG GCGCCAAGTA | 1320 |
| 20 | GCAAAAATTG GTACTGCCTT TGAAGCGACA AATCATGAAT CAACATTACA ATACGTATTC | 1380 |
| | GGTGAAAAGT TTAGTAAAGT CTTTGATTAT ATTTTAATCT TCTTCTTATT TGGTATAGCT | 1440 |
| 25 | GTAACCATGC LAGCTGGTGC AGGCGCAACA TTTGAAGAAA GTTATAACAT ACCTACATGG | 1500 |
| | CTAGGTGCTT TAATTATGAC ATTAGCGATT TATATTACGT TGCKATTAGA CTTTAATAAA | 1560 |
| | ATAGTACGTG CACTAGGTAT CGTTACACCA TTTTTAATTG TTTTAGTTGT ATTAATCGCT | 1620 |
| 30 | GGCGTTTATT LATTTAAAGG TCATGLTTCA TTAGCAGAAG TTAACCAAGT AGTGCCLGAA | 1680 |
| • | GCAAGTATTT GGAAGGGAAT CTGGTTTGGT ACAATATATG GTGGATTAGC TTTTTCTGTA | 1740 |
| | GGTTTTAGTA CCATCGTAGC AATCNGTGGG GATACTGAAA AGCGTACAGT GTCAGGTGCA | 1800 |
| 35 | GGCGCGATGT ATGGTGGTAT TATCTATACT GTATTACTAG CATTGATCAA CTTTGCATTG | 1860 |
| | CAAGTGAATA TCCAACTATT AAAAATGCCT CAATTCCTAC ATTGACGTTA GCAAATAATA | 1920 |
| 40 | TCCATCCTTT AATAGCAACA GTGKTATCTG TTATTATGCT GGCGGKTATG TATAATACTA | 1980 |
| | TTCTAGGACT AATGTATTCA TTTGCAGCAC GTTTTACAGA ACCATACAGT AAAAATTATC | 2040 |
| | ATATCTTTAT TATTATAATG ATGGTAGCAG GTTATTTATT AAGTTnCGTA GGATTTGCTG | 2100 |
| 45 | AATTAATTAA TAAGTTATAT ACNATTTATG GGATATGTAG GCTTATTNTA TTGTAGTAGC | 2160 |
| | TGTAATTATN AAATATTTCC AAACGTAAAA ATGGCGGATA AAAAACATAT TGCTTTAATA | 2220 |
| | TCATATGGAG GGGATATCCG AAACTTTACA ATTTGAATCA CTTTGGT | 2267 |
| 50 | (2) INFORMATION FOR SEC ID NO. 219. | |

(2) INFORMATION FOR SEQ ID NO: 219:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6336 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
- 55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 219:

| 5 | GTCAATGTAA | CCTAATAGTT | TATGTCTATC | TTGTGTACCA | ACTACTACAT | CGACACCAGG | 60 |
|----|------------|------------|------------|------------|------------|------------|------|
| | AATTTCCATA | ATTTCAGCTG | ATGAAGTTTG | CGCATAACAA | CCTGTTACAC | AGATTACAGC | 120 |
| 0 | ATCAGGATTT | TGTCTTATTG | CACGTCTAAT | TATTTGACGA | CTTTTTTAT | CACCCGTATT | 180 |
| | CGTTACTGTA | CAAGTATTAA | TAACAAATAC | ATCAGCATTC | GCTTCAAAGT | CAACGCGCTC | 240 |
| | ATAGTTTGCT | TCTTTAAATA | ATTGCCAGAT | TGCTTCAGTT | TCATAATGGT | TTACTTTACA | 300 |
| 5 | ACCTAATGTG | TGAACGCAAC | TGTTGACATA | AATATTCACC | CCATTAATTC | TTTTTCATAA | 360 |
| | CTTATTGCAC | TTAACGCATA | CAATGGCGCA | GTTTCTGCCC | GTAAAATTCT | CGGCCCAAGA | 420 |
| | CCAACAACTG | TACTAGTATT | ACTAAATAAT | GAAATTTCAT | TTTCTGACAA | ACCACCCTCA | 480 |
| 0 | GGaCCAAAAA | TCATCAACAC | TTTATCCTGA | GCATTGAATT | GTTGTAAAGT | TTGCTTGAAA | 540 |
| | TTGCTTAACT | CACCATCTTT | TGCTTCCTCT | TCATATGCAA | TAAGAATATA | GTCATAATTA | 600 |
| r | TCAATAGTAT | CACAAATTAA | TTTTAAATTC | GACTCGAATT | GAATAGATGG | AATCACTAAA | 660 |
| 5 | CGATAGCTTT | GTTCAGCAGC | TTCTTTAATT | ATTTTTTGCC | AACGCTCTAT | CTTTTTGGCA | 720 |
| | ACTITIGCCT | CGTTTAATTT | AACAATTGAA | CGTTCCATGC | TCACAGCTAT | AAATGATGAA | 780 |
| 0 | GCACCCAATT | CAGTAGCTTT | TTGTAGCAAC | CACTCATATT | TGTCAGCTTT | GATTAGTCCA | 840 |
| | CTGCAAATCG | TAACATCAAC | TGGCAATTCT | GTATTAATAT | TTTGTTTTTC | TTTTAAATCA | 900 |
| | ACTTCAATTT | TATCACTTGT | TATGTCAGCA | ATTTCACATA | AATAAACTGT | TTGATCATTA | 960 |
| 5 | AAAGTTAAAA | TAATTTTACT | ACCAACATCA | TATCTCATTA | CATTTGTTAT | ATGATGAATA | 1026 |
| | TCTTCTTTTT | TIGTAATAAA | AAAACGCTGA | CTTACATCAG | CGTTTTGGnT | CTATGAAATA | 1080 |
| | ACGITGCACA | TTATTCACTC | ACTTTCTGGC | CAACAAGACA | AACCCAACCG | TTGTCATGTT | 1140 |
| 0 | GTTCTGAAAT | AATTTTAAAA | CCTACACGCT | CCATATGTGA | CTGTATACCT | TCATACTTCT | 1200 |
| | CTTTTATAAT | ACCAGAAGTA | ATAAAATAAC | CGCCTTCATT | TAGAGTATTA | TAAGCATCTT | 126 |
| 5 | CAATCATTTC | ATCAATAATA | TGCGCTAAAA | TATTTGCTAT | TACAATATCA | AATTTTTCTG | 132 |
| | TTTCGTCTTT | CAATAAGTTA | CCTGGAACAG | CTTCAATTAA | CGTTTCACAA | TGATTTCTTC | 138 |
| | TGAAGTTTTC | TTTAGCTACA | CTCACTGCCA | TTTCATCAAT | ATCCAACGCT | TTAATACGTT | 144 |
| io | TTACACCGAT | TAGATGACTT | GCAATACTTA | ATATACCTGA | GCCAGTACCA | ACATCAATTA | 150 |
| | CTGAATGCTG | TGGCAATACA | TATGTTTCTA | TTGCCTTCAA | ACACATACTT | GTAGTCGGAT | 156 |
| | GATCACCTGT | TCCAAAAGCC | ATACCTGGGT | CGAGCTCAAT | GCAAAGCTCT | TCATCCGCTT | 162 |

| | GGAAATAGT. | 1 TITCCATTCA | TITTCCCAA | T CCGTCTCTG | C AATAATTTGO | TCACTGAATT | 174 |
|----|------------|--------------|------------|--------------|--------------|------------|-------|
| _ | GAACGTTAT | G TTGATCAAGT | TCATCTAAA | T TTAATAACTO | ATCTTTAATT | TGCTGTCGCA | 180 |
| 5 | ACTTATCATO | C ATAAGTCATT | TCATTAAAA | T AGGCTTTCA | TCTTACTCC | TTATCTGGAT | 186 |
| | AATCCTCTT | TTTCAAAGCG | TAAATTTCAC | CGTATTTATO | TTCTGGTTGG | TTAATTAAAT | 192 |
| 10 | CATCTGAATO | TTCTATCACG | ACACCATTTO | ATCCATGATT | TTCAAGTATA | TTGGTAGCCA | 1980 |
| | ATTCTACTG | TTCATGATTA | ATAATAATTO | AAAGCTCTGT | CCAGTTCATA | CTTTATTCTC | 2040 |
| | CCTTAAAGAA | TCTTTTTGCT | CTATCTTTAA | AATTCGAAGG | TTGTTCATTA | ATTTCTTCAC | 2100 |
| 15 | CATTTAATTG | GGCAAATTCT | TTCATTAGTT | CITTITGTCT | ATCTGTTAAT | TTAGTAGGCG | 2160 |
| | TTACTACTTT | AATATCAACA | TATAAATCTC | CGTATCCATA | GCCATGAACA | TTTTTTATAC | 2220 |
| | CCTTTTCTTT | TAAGCGGAAT | TGCTTACCTG | TTTGTGTACC | AGCAGGGATT | GTTAACATAA | 2280 |
| 20 | CTTCATTATT | TAATGTTGGT | ATTTTTATTT | CATCGCCTAA | AGCTGCTTGT | GGGAAGCTAA | 2340 |
| | CATTTAATTT | GTAATAAATA | TCATCACCAT | CACGTTTAAA | TGTTTCAGAT | GGTTTAACTC | 2400 |
| 25 | TAAATACTAC | GTATAAATCA | CCAGCAGGTC | CTCCATTCAC | GCCTGGAGAG | CCTTCACCAG | 2460 |
| | СТААТСТААТ | TTGTTGTTCA | TTGTCGACAC | CTTCAGGTAC | TTTCACTTCT | AATTTAACTG | 2520 |
| | TTTTATTTTC | AGTACCTTTT | CCGTGACATG | TTGGACAAGC | TTCTTCAAAT | TCTTGACCAC | 2580 |
| 30 | TTCCATTACA | TTTAGGACAA | ACTTGTTCAG | TACGAACTCT | ACCTAAAATT | GTGTTTTGTT | 2,640 |
| | CTACAGCTAC | ATGACCAGCG | CCATTACAGT | AACTACAAGT | CTTTTTACTT | GTTCCAGGCT | 2700 |
| | TTGCACCATC | ACCATGACAT | GTTTCGCATG | TTACATCTTT | ACGGATTGAA | ATTTCTTTTG | 2760 |
| 15 | TTGTACCAAA | TACCGCTTCT | TCAAATGTTA | ATGTCATTGT | ATACTGAAGA | TCATCACCTT | 2820 |
| | TTTGCGGTGC | ATTTGGATCT | CTTTGTCTGC | CGCCACCGAA | GAAAGAGCTA | AAGATATCTT | 2880 |
| _ | CAAABCCGCC | GCCACCGAAG | CCACTAAAAC | CGCCAAAGTC | AGAGCCATTG | AATCCTTGTC | 2940 |
| 0 | CACCAAAACC | TTGTGGACCA | TCATGTCCAA | ATTGATCATA | GcTTGCGCGT | TTATTATCAT | 3000 |
| | CACTTAAAAC | TTCATAGGCT | TCAGAAATTT | CTTTAAACTT | TTCATCTGCA | CCTTCTTCTT | 3060 |
| 5 | TGTTAATATC | TGGATGATAT | TTTTTCGAAA | GCTTTCGATA | CGCTTTTTTG | ATTTCATCTT | 3120 |
| | TTGAAGCATC | CTTACTAATG | CCTAAAACTT | CATAATAATC | TCTTTTGGCC | ACAGCTATCT | 3180 |
| | CTCCTTTTCT | TAATTAACTC | ATATAGTTTA | ACGTAATATG | TCATACTATC | САААТААААА | 3240 |
| 0 | GCCAAAGCCA | ATGTTCTATT | GACTTTGACT | TTTCAGATCA | TGACAACATT | CTAATTGTAT | 3300 |
| | TGTTTAATTA | TTTTTTGTCG | TCGTCTTTTA | CTTCTTTAAA | TTCAGCATCT | TCTACAGTAC | 3360 |
| | TATCATTGTT | TTGACCAGCA | TTAGCACCTT | GTGCTTGTTG | TTGCTGTTGA | GCCGCTTGCT | 3420 |

970

| | TATCTTCTAT | ATCTTGACCT | TCTAAAGCAG | TTTTAAGAGC | GTCTTTTTTC | TCTTCAGCAG | 3540 |
|-----------|------------|------------|------------|------------|------------|------------|------|
| _ | ATTTTTTATC | TTCTTCACCG | ATATTTTCGC | CTAAATCAGT | TAAAGTTTTT | TCAACTTGGA | 3600 |
| 5 | ATACTAGACT | GTCAGCTTCG | TTTCTTAAGT | CTACTTCTTC | ACGACGTTTT | TTATCTGCTT | 3660 |
| | CAGCGTTAAC | TTCAGCATCT | TTTACCATAC | GGTCGATTTC | TTCGTCTGAT | AATGAAGAAC | 3720 |
| 10 | TTGATTGAAT | TGTAATTCTT | TGTTCTTTAT | TTGTACCTAA | GTCTTTTGCA | GTTACATTTA | 3780 |
| | CAATACCGTT | TTTATCGATA | TCAAACGTTA | CTTCAATTTG | AGGTTTACCA | CGTTCAGCTG | 3840 |
| | GTGGAATATC | AGTCAATTGG | AATCTACCAA | GTGTTTTATT | ATCCGCAGCC | ATTGGACGTT | 3900 |
| 15 | CACCTTGTAA | TACGTGTACA | TCTACTGATG | GTTGATTATC | TACTGCTGTT | GAATAGATTT | 3960 |
| | GAGATTTAGA | TGTAGGAATC | GTAGTGTTAC | GTTCAATTAA | CGTATTCATA | CGTCCACCTA | 4020 |
| | AAATTTCAAT | ACCTAAAGAT | AGTGGTGTTA | CGTCTAATAA | TACTACGTCT | TTAACGTCAC | 4080 |
| 20 | CTGTGATAAC | GCCACCTTGG | ATTGCAGCTC | CCATTGCCAC | TACTTCGTCC | GGGTTTACTC | 4140 |
| | CTTTGTTAGG | CTCTTTACCG | ATTTCTTTTT | TGACAGCTTC | TTGTACTGCT | GGAATACGAG | 4200 |
| 25 | TTGATCCACC | AACTAAGATA | ACTTCATCGA | TATCTGAGTT | TGTTAAGCCA | GCGTCTTTCA | 4260 |
| | TTGCTTGGCG | TGTAGGTTCC | ATTGTTCTTC | TAATTAATGA | ATCTGATAAT | TCTTCAAATT | 4320 |
| | TAGAACGAGT | TAAGTTTACT | TCTAAGTGTA | ATGGACCGTT | TTCACCAGCT | GAGATAAATG | 4380 |
| 30 | GTAATGAGAT | TTGAGTTTGT | GATACACCTG | ATAAGTCTTT | TTTAGCTTTT | TCAGCAGCAT | 4440 |
| | CTTTCAAACG | TTGTAATGCC | ATTTTATCTT | GAGATAAGTC | TACGCCATTT | TCTTTTTGA | 4500 |
| | ATTCTGCAAC | TAGGTAGTCA | ATAATTACTT | GGTCAAAATC | ATCACCGCCA | AGTTTGTTGT | 4560 |
| 35 | CACCGGCTGT | TGATAGTACT | TCGAATACAC | CGTCACCTAA | TTCTAGGATA | GATACGTCAA | 4620 |
| | ATGTACCGCC | ACCTAAGTCA | AAAACAAGAA | CTTTTTCATC | TTTATCAGTT | TTGTCTAAAC | 4680 |
| 40 | CATATGCTAA | TGCTGCAGCT | GTTGGTTCAT | TAATGATACG | CTCAACTTCT | AAACCAGCAA | 4740 |
| 40 | TTTTACCAGC | ATCTTTAGTT | GCTTGACGTT | CAGCATCGTT | AAAGTATGCA | GGTACTGTAA | 4800 |
| | TTACAGCTTT | GTCAACTTTC | TCACCTAAAA | TAGTTTCAGC | TGTATTTTT | AAGTTTTGTA | 4860 |
| 45 | AAATCATAGC | TGAGATTTCT | TGTGGTGTGT | ATGATTTACC | TTCAATATCT | ACTTTATAAT | 4920 |
| | CAGTACCCAT | ATGACGTTTA | ATAGATTGAA | CAGTGTTTGG | GTTTGTAATA | GCTTGACGTT | 4980 |
| | TTGCTACTTC | accaacttga | GTTTCTCCAT | TTTTGAAAGC | TACAACAGAT | GGTGTTGTAC | 5040 |
| 50 | GTGAACCTTC | AGGGTTTTGA | ATTACTTTTG | GCTCATCGCC | TTCTAATACT | GTNACACATG | 5100 |
| | AATTTGTTGT | ACCTAAGTCT | ATACCAATAA | TTTTACTCAT | AATAAAATTO | CTCCATTTAA | 516 |
| | TCATTAAATT | AATTTAATTT | TAAACAATGI | CTTTTCGCCA | AATTTAAGTT | ATTGGTTTAC | 522 |

| AATCAAACTI TCATGCACCA TITGTACACC TITTTGAAGA GATTTAAAAG TCTCATCATC ACCITCAATI TGAAGTGCAC GITCTATATI GTCTATTGCT GGTAAAATAT CTGTTAACAC ACGITGTGCT TGATATGTTT TGTTTATTC ATTTTCTTTT TGAATTCTAC GCTTATAATT TTCAAACTCA GCGTAGAGCC TTAAATATTT CTCTTCGTTT TCATCTGCTA ATTGTTGAAG TTCATTAATT TTTTGATCTT TTGGATCTAT TTCTTCAATA ACATTCTCGT CAGACGTTTC AATATTTGT TTGACGTTTG TTTCTTCAAC TGTTGATTCA GCTTGTTCAAC CTGAATCATC GTCTTTATTT GTCATTTTCT GTCCTCCAAT ACTTTCTAAT CCATCATTAC CAAATTCTAT GTCATTAATT TGAACATTT TGATAATGCA TAGCTGTAGG TCCAATCACA GCGATTTGAC CTTTTAACGT TTCATCAAAA TGATATTGAC TTGTTACAAT TGAAATATCA CTTAAGCTGT CATCAATTC ATTACCAATT TTTACATTAA TATTTGGTGA AGATTATCT TGTAATAATT CATCAATTC ATTACCAATT TTAACATTA TATTTGGTGA AGATTATAC CTTTGATTAG CATCAATTC ATTAGATTAT TAACCTTC CACCCATATA AATGCTATTA CTTTGATTAG AAATATGATT ATTCATCGTA TTTAACCATT TATTGATAAA AATTTCTTCC TGCTCTGATT GAACAAAAGA GACAATATCA TCTTGTAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAA CANAATTCAA TATCATCAAT TCTTGTAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAA CANAATTCAA TATCATCAAT TTTAACAATT TATTGATAAAA AATTTCTTCC TGCTCTGATT GAACAAAAGA GACAATATCA TCTTGTAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAA CANAATTCAA TATCATCAATA TCTTGTAAAAT TCTGATTAAAA CTCAGTTAGT TTGTTTGTAAA CANAATTCAA TATCATCAATA TCTTGTAAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAAA CANAATTCAA TATCATCAATA TCTTGTAAAAT TCTGATTAAAA CTCAGTTAGT TTGTTTGTAAA CANAATTCAA TATCAATATAA TCTGTTTAAAATTAAAATTAAAATTAAAATTAAAATTAAAATTAAAA | |
|--|-----|
| ACCITICAATT TGAAGTGCAC GITCTATATT GICTATTGCT GGTAAAATAT CTGTTAACAC ACGTTGTGCT TGATATGTTT TGTTTATTTC ATTTTCTTTT TGAATTCTAC GCTTATAATT TTCAAACTCA GCGTAGAGCC TTAAATATTT CTCTTCGTTT TCATCTGCTA ATTGTTGAAG TTCATTAATT TTTTGATCTT TTGGATCTAT TTCTTCAATA ACATTCTCGT CAGACGTTTC TTCATTACT TCATCTGTA AATGACCTTT ACTTTCTAA GCTTGTTCAA CTGAATCATC AATATTTTGT TTGACGTTTG TTTCTTCAAC TGTTGATCA GCTGTTTTT CAACTGATTC GTCTTTATTT GTCATTTTCT GTCCTCCAAT ACTTTCTAAT CCATCATTAC CAAATTCTAT TTAATAATTG AATGACATTT TGATAATGCA TAGCTGTAGG TCCAATCACA GCGATTTGAC CTTTTAACGT TTCATCAAAA TGATATTGAC TTGTTACAAT TGAAATATCA CTTAAGCTGT CATCAATTC ATTACCAATT TTTACATTAA TATTTGGTGA AGATATATCT TGTAATAATT CATCAATCT ATTTGATTCT ATATATTGTA GAATGGGCTG AATTGAAGAT ACATTACTTT CATCAATGC ATCAATAAGT TTAACCTTTC CACCCATATA AATGCTATTA CTTTGATTAG AAATATGAT ATTCATCGTA TTTAACATT TATTGATAAA AATTTCTTCC TGCTCTGATT GAACAAAAGA GACAATATCA TCTTGTAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAA CAAAATTGA TATTGTATTT AGTTTGTCAT TATTAA (2) INFORMATION FOR SEQ ID NO: 220: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | 400 |
| ACGITGECT TGATATGITT TGTTTATTC ATTITCTITT TGAATTCIAC GCTTATAATT TTCAAACTCA GCGTAGAGCC TTAAATATT CTCTTCGTTT TCATCTGCTA ATTGTTGAAG TTCATTAATT TTTTGATCTT TTGGATCTAT TTCTTCAATA ACATTCTCGT CAGACGTTTC TTCATTGCT TCATCTTGTA AATGACCTTT ACTITCTTCA GCTTGTTCAA CTGAATCATC AATATTTGT TTGACGTTTG TTTCTTCAAC TGTTGATTCA GCTGTTTTTT CAACTGATTC GTCTTTATTT GTCATTTTCT GTCCTCCAAT ACTTCTAAT CCAACTCATTAC CAAATTCTAT TTAATAATTG AATGACATTT TGATAATGCA TAGCTGTAGG TCCAATCACA GCGATTTGAC CTTTTAACGT TTCATCAAAA TGATATTGAC TTGTTACAAT TGAAATATCA CTTAAGCTGT CATCAATTTC ATTACCAATT TTTACATTAA TATTTGGTGA AGATATATCT TGTAATAATT CTGCAATTCT ATTTGATTCT ATATATTGTA GAATGGGCTG AATTGAAGAT ACATTACTTT CATCAATGC ATCAATAAGT TTAACCATT TATTGATAAA AATGCTATTA CTTTGATTAG AAAATTGAAT ATTCATCGTA TTTAACCATT TATTGATAAA AATTCTTCC TGCTCTGATT GAACAAAAGA GACAATATCA TCTTGTAAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAA CAAAATTGA TATTGTATTT AGTTTGCAT TATTAA (2) INFORMATION FOR SEQ ID NO: 220: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | 460 |
| TTCAAACTCA GCGTAGAGCC TTAAATATT CTCTTCGTTT TCATCTGCTA ATTGTTGAAG TTCATTAATT TTTTGATCTT TTGGATCTAT TTCTTCAATA ACATTCTCGT CAGACGTTTC TTCTATTGCT TCATCTTGTA AATGACCTTT ACTTTCTTCA GCTTGTTCAA CTGAATCATC AATATTTTGT TTGACGTTTG TTTCTTCAAC TGTTGATTCA GTGTTTTTTT CAACTGATTC GTCTTTATTT GTCATTTCT GTCCTCCAAT ACTTTCTAAT CCATCATTAC CAAAATTCTAT TTAATAATTG AATGACATTT TGATAATGCA TAGCTGTAGG TCCAATCACA GCGATTTGAC CTTTTAACGT TTCATCAAAA TGATATTGAC TTGTTACAAT TGAAATATCA CTTAAGCTGT CATCAATTC ATTACCAATT TTTACATTAA TATTTGGTGA AGATATATCT TGTAATAATT CTGCAATTCT ATTTGATTCT ATATATTGTA GAATGGCTG AATTGAAGAT ACATTACTTT CATTCAATGC ATCAATAAGT TTAACCATTT CACCCATATA AATGCTATTA CTTTGATTAG AAAATATGATT ATTCATCGTA TTTAACAATT TATTGATAAA AATTTCTTC TGCTCTGATT GAACAAAAGA GACAATATCA TCTTGTAAAT TCTGATTAAA CTCAGGTTAGT TTGTTTGTAA CAAAATTTGA TATTGTATTT AGTTTGTCAT TATTAA (2) INFORMATION FOR SEQ ID NO: 220: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | 520 |
| TTCATTAATT TTTTGATCTT TTGGATCTAT TTCTTCAATA ACATTCTCGT CAGACGTTTC TTCTATTGCT TCATCTTGTA AATGACCTTT ACTTTCTTCA GCTTGTTCAA CTGAATCATC AATATTTTGT TTGACGTTTG TTTCTTCAAC TGTTGATTCA GTGTTTTTTT CAACTGATTC GTCTTTATTT GTCATTTCT GTCCTCCAAT ACTTTCTAAT CCATCATTAC CAAAATTCTAT TTAATAATTG AATGACATTT TGATAATGCA TAGCTGTAGG TCCAATCACA GCGATTTGAC CTTTTAACGT TTCATCAAAA TGATATTGAC TTGTTACAAT TGAAAATATCA CTTAAGCTGT CATCAAATTC ATTACCAATT TTTACATTAA TATTTGGTGA AGATATATCT TGTAATAATT CATCAAATCT ATTTGATTCT ATATATTGTA GAATGGGCTG AATTGAAGAT ACATTACTTT CATTCAATGC ATCAATAAGT TTAACCATT CACCCATATA AATGCTATTA CTTTGATTAG AAATATGATT ATTCATCGTA TTTAACAATT TATTGATAAA AATTCTTCC TGCTCTGATT GAACAAAAGA GACAATATCA TCTTGTAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAA CAAAATTTGA TATTGTATTT AGTTTGTCAT TATTAA 63 (2) INFORMATION FOR SEQ ID NO: 220: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | 580 |
| TICTATIGCT TCATCTIGTA AATGACCTIT ACTITCTICA GCTTGTTCAA CTGAATCATC AATATTTIGT TIGACGTTIG TITCTTCAAC TGTTGATTCA GTGTTTTTT CAACTGATTC GTCTTTATTT GTCATTTTCT GTCCTCCAAT ACTITCTAAT CCATCATTAC CAAATTCTAT TTAATAATTG AATGACATTT TGATAATGCA TAGCTGTAGG TCCAATCACA GCGATTTGAC CTTTTAACGT TTCATCAAAA TGATATTGAC TTGTTACAAT TGAAATATCA CTTAAGCTGT CATCAATTC ATTACCAATT TTTACATTAA TATTTGGTGA AGATATACT TGTAATAATT CATCAATGC ATCAATAAGT TTAACCTTTC CACCCATATA AATGCTATTA CTTTGATTAG AAATATGATT ATTCATCGTA TTTAACAATT TATTGATAAA AATTTCTTCC TGCTCTGATT GAACAAAAGA GACAATATCA TCTTGTAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAA CAAAATTTGA TATTGTATTT AGTTTGTCAT TATTAA 63 (2) INFORMATION FOR SEQ ID NO: 220: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | 540 |
| AATATTTET TEGACETTE TTTCTTCAAC TETTGATTCA GETTGTTTT CAACTGATTC GTCTTTATTT GTCATTTTCT GTCCTCCAAT ACTTTCTAAT CCATCATTAC CAAATTCTAT SE GTCTTTATTT GTCATTTTCT GTCCTCCAAT ACTTTCTAAT CCATCATTAC CAAATTCTAT TTAATAATTG AATGACATTT TGATAATGCA TAGCTGTAGG TCCAATCACA GCGATTTGAC CTTTTAACGT TTCATCAAAA TGATATTGAC TTGTTACAAT TGAAATATCA CTTAAGCTGT CATCAATTTC ATTACCAATT TTTACATTAA TATTTGGTGA AGATATACTT TGTAATAATT CTGCAATTCT ATTTGATTCT ATATATTGTA GAATGGGCTG AATTGAAGAT ACATTACTTT CATTCAATGC ATCAATAAGT TTAACCTTTC CACCCATATA AATGCTATTA CTTTGATTAG AAATATGATT ATTCATCGTA TTTAACAATT TATTGATAAA AATTTCTTCC TGCTCTGATT GAACAAAAGA GACAATATCA TCTTGTAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAA CAAAATTTGA TATTGTATTT AGTTTGTCAT TATTAA (2) INFORMATION FOR SEQ ID NO: 220: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (YI) SEQUENCE DESCRIPTION, SEQ ID NO: 220. | 700 |
| GTCTTTATTT GTCATTTTCT GTCCTCCAAT ACTTTCTAAT CCATCATTAC CAAATTCTAT TTAATAATTG AATGACATTT TGATAATGCA TAGCTGTAGG TCCAATCACA GCGATTTGAC CTTTTAACGT TTCATCAAAA TGATATTGAC TTGTTACAAT TGAAATATCA CTTAAGCTGT CATCAATTTC ATTACCAATT TTTACATTAA TATTTGGTGA AGATATATCT TGTAATAATT CTGCAATTCT ATTTGATTCT ATATATTGTA GAATGGGCTG AATTGAAGAT ACATTACTTT CATTCAATGC ATCAATAAGT TTAACCTTTC CACCCATATA AATGCTATTA CTTTGATTAG AAATATGATT ATTCATCGTA TTTAACCATT TATTGATAAA AATTTCTTCC TGCTCTGATT GAACAAAAGA GACAATATCA TCTTGTAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAA CAAAATTTGA TATTGTATTT AGTTTGTCAT TATTAA 63 (2) INFORMATION FOR SEQ ID NO: 220: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | 760 |
| TTAATAATTG AATGACATTT TGATAATGCA TAGCTGTAGG TCCAATCACA GCGATTTGAC CTTTTAACGT TTCATCAAAA TGATATTGAC TTGTTACAAT TGAAATATCA CTTAAGCTGT CATCAATTTC ATTACCAATT TTTACATTAA TATTTGGTGA AGATATATCT TGTAATAATT CTGCAATTCT ATTTGATTCT ATATATTGTA GAATGGGCTG AATTGAAGAT ACATTACTTT CATTCAATGC ATCAATAAGT TTAACCTTTC CACCCATATA AATGCTATTA CTTTGATTAG AAATATGATT ATTCATCGTA TTTAACAATT TATTGATAAA AATTTCTTCC TGCTCTGATT GAACAAAAGA GACAATATCA TCTTGTAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAA CAAAATTTGA TATTGTATTT AGTTTGTCAT TATTAA 63 (2) INFORMATION FOR SEQ ID NO: 220: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | 320 |
| TTAATAATTG AATGACATTT TGATAATGCA TAGCTGTAGG TCCAATCACA GCGATTTGAC CTTTTAACGT TTCATCAAAA TGATATTGAC TTGTTACAAT TGAAATATCA CTTAAGCTGT CATCAATTTC ATTACCAATT TTTACATTAA TATTTGGTGA AGATATATCT TGTAATAATT CTGCAATTCT ATTTGATTCT ATATATTGTA GAATGGGCTG AATTGAAGAT ACATTACTTT CATTCAATGC ATCAATAAGT TTAACCTTTC CACCCATATA AATGCTATTA CTTTGATTAG AAATATGATT ATTCATCGTA TTTAACAATT TATTGATAAA AATTTCTTCC TGCTCTGATT GAACAAAAGA GACAATATCA TCTTGTAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAA CAAAATTTGA TATTGTATTT AGTTTGTCAT TATTAA (2) INFORMATION FOR SEQ ID NO: 220: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | 80 |
| CATCAATTC ATTACCAATT TTTACATTAA TATTTGGTGA AGATATATCT TGTAATAATT CTGCAATTCT ATTTGATTCT ATATATTGTA GAATGGGCTG AATTGAAGAT ACATTACTTT CATTCAATGC ATCAATAAGT TTAACCTTTC CACCCATATA AATGCTATTA CTTTGATTAG AAATATGATT ATTCATCGTA TTTAACAATT TATTGATAAA AATTTCTTCC TGCTCTGATT GAACAAAAGA GACAATATCA TCTTGTAAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAA CAAAATTTGA TATTGTATTT AGTTTGTCAT TATTAA 63 (2) INFORMATION FOR SEQ ID NO: 220: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | 940 |
| CTGCAATTCT ATTTGATTCT ATATATTGTA GAATGGGCTG AATTGAAGAT ACATTACTTT CATTCAATGC ATCAATAAGT TTAACCTTTC CACCCATATA AATGCTATTA CTTTGATTAG AAATATGATT ATTCATCGTA TTTAACAATT TATTGATAAA AATTTCTTCC TGCTCTGATT GAACAAAAGA GACAATATCA TCTTGTAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAA CAAAATTTGA TATTGTATTT AGTTTGTCAT TATTAA (2) INFORMATION FOR SEQ ID NO: 220: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | 000 |
| CATTCAATGC ATCAATAAGT TTAACCTTTC CACCCATATA AATGCTATTA CTTTGATTAG AAATATGATT ATTCATCGTA TTTAACAATT TATTGATAAA AATTTCTTCC TGCTCTGATT GAACAAAAGA GACAATATCA TCTTGTAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAA CAAAATTTGA TATTGTATTT AGTTTGTCAT TATTAA (2) INFORMATION FOR SEQ ID NO: 220: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | 60 |
| AAATATGATT ATTCATCGTA TTTAACAATT TATTGATAAA AATTTCTTCC TGCTCTGATT GAACAAAAGA GACAATATCA TCTTGTAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAA CAAAATTTGA TATTGTATTT AGTTTGTCAT TATTAA (2) INFORMATION FOR SEQ ID NO: 220: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | .20 |
| GAACAAAAGA GACAATATCA TCTTGTAAAT TCTGATTAAA CTCAGTTAGT TTGTTTGTAA CAAAATTTGA TATTGTATTT AGTTTGTCAT TATTAA 63 (2) INFORMATION FOR SEQ ID NO: 220: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | .80 |
| CAAAATTTGA TATTGTATTT AGTTTGTCAT TATTAA (2) INFORMATION FOR SEQ ID NO: 220: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | 40 |
| (2) INFORMATION FOR SEQ ID NO: 220: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | 00 |
| (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | 36 |
| (A) LENGTH: 13059 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (Xi) SECUENCE DESCRIPTION: SEC. ID NO. 220. | |
| (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (Xi) SECUENCE DESCRIPTION: SEC. ID No. 220. | |
| (D) TOPOLOGY: linear | |
| (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 220: | |
| (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 220: | |
| | |
| TTCATGATTA TTATCTGTTG TAGACACTGC TGGATCTTCC GATGTATCTT TCGATGCATC | 60 |
| TTTCGATTTG TGTATTTGCT GATTCAAATG GTCTAGGTCT TCTAACGCCT TATTTACCAT 1 | 20 |
| TGCTTCATCA TTTTTATCAT CTTTTTCTCC ATGTTTTGTT GTAGCCGTTT GTGACATATC | 80 |
| ATTTTCATT GCATTAAGAT CGTCCTCGCC ACTTTGTTGA CCCCTATCAA CATTTGAAGA 2 | 40 |
| AACCTCATTT AAATCTTTAA GCAATTGATC TAATTTACTG TCTATATCAC TTTGACCGTT 3 | 00 |

| | TTCATCTATT | TGCGATGCTG | TFFTCGCTTC | ATTTAGTTGT | GCTTTATAAT | GITCTTTAGA | 420 |
|----|------------|------------|------------|------------|-------------|------------|------|
| 5 | TGAAGCCGAT | AACTGTTTTA | ATTGCTCAAT | TTGACGAATT | GCCTTGTCAA | CTTTGTCTAA | 480 |
| | TAAATCTTGC | TTAGATAATA | TCTCTTTTGT | AATTTCAGTA | TCCTTTTCAG | ATGCAGCTTG | 540 |
| | GGCATCGTAC | GGCAAGATAT | TCGTTAAAAT | GATACTTGTC | GCCATCATTG | TCGAACACGA | 600 |
| 10 | TAACTTTACA | TATAATTGAA | ACGGTTTCCC | TCGATATTTA | GCCATCAACA | TACTCCTTCC | 660 |
| | TCACTTACTT | CCTTCAAAGA | ATTACATACT | ATTATATACC | TGTTTACAAG | AAATTTACAC | 720 |
| | TTATCTATCT | AGTTATTGTT | GTTAGTAATT | ATCTACTTAT | TACTTAGCTT | ATATTTAAGT | 780 |
| 15 | AAACAAAACA | AGCATGACGT | AATATCATAT | TGTCCATGTC | GCTAACATCA | TATTACGTCA | 840 |
| ٠ | AATCTTTTAT | ATTAAATGAT | GTTTTATTTT | AGACTGCTTT | TTCCTTTTAG | CTTTCGAGCG | 900 |
| | CCTGTTTAAA | AACTTGCTCG | AATTGTTCAC | GCGAGATTTC | GTGTGCATGT | GCTTTTTGTG | 960 |
| 20 | CTAATAAAGC | ATCTCGAAAC | TGTTGTTGAT | CTTTCAAACT | TTCTAACATT | TGTATTAATT | 1020 |
| | GGTCTTTACT | TTCCATTGTT | ATCTCATCAT | TATGCTCAAA | TAAGTGCTCT | GATAATGTTA | 1080 |
| 25 | CTTTAGCATG | GTGTGCGGTT | TGACGATAAC | CTAAAATCAA | CAACTCATAG | TCAAACGCTT | 1140 |
| | GTTCCACCGC | ATTTAAAATT | TCATTACCCT | CATTGATATC | AAGATAAATA | TCACATAACT | 1200 |
| | GGTATAGTTC | ATTTACCCTG | TCAATATTAA | TAGATGGGTA | TAAATGCACA | TTAGCATATT | 1260 |
| 30 | GATCAAGTTG | CATTAGCTTA | TCAGACATCT | CTGTAATAGC | AGCGATGTGn | AACTTAAAAT | 1320 |
| | CTGGTAAAGt | TYCAACCAAT | ACCTTGATGT | TACGAatTGa | TCCgAGTTAG | TTAATATTAC | 1380 |
| | AATTTCTTTA | GTATATCTAT | TACGACTACG | ATAGTTATAT | AGATATCCGC | CTTGTAAAAT | 1440 |
| 35 | ACGAGATTGA | ACCTTTGCGT | CTGCTATATT | GAGCATCGTT | TCATATTCGT | TTTTATCTGG | 1500 |
| | ATAATAATA | TTACAATGTC | GTTTCATATC | ACCTTTACAC | ATCAATTGCA | TATTTCCCGG | 1560 |
| 40 | GACATTACCA | TTACAGTGTT | CTTGCCATAC | CAAAACATCA | CTACCTTTTG | ATGGCAAATT | 1620 |
| | ATATAACACT | GAAAATGGTA | GGGCTAGTGA | GTTAATAACG | AAATGATGTT | CCGTAATTTC | 1680 |
| | AAGTTGCTTG | ATAAAAAATA | ATGCGAATGC | GAGCTTTGAA | GGGAAAAAGT | AAGACTTCCC | 1740 |
| 45 | TTGCCAATCC | AATATGACAT | CAGATGTTAC | AAAATTTTCA | TAAATCACTT | CTTTACCTTC | 1800 |
| | TGCTGTCATA | TATTTCTTCA | AGATCGCTTT | ACGATTTAAA | TCGTAAACAG | TTTGTGCAAA | 1860 |
| | TTTAATACCA | TTCTTAGAAT | AATAATCGAC | AAATCGGACA | CGTTGTTGGT | CATCAAACCA | 1920 |
| 50 | TTCGACACGA | CTAACAATTC | TAGGGCGCTC | TCCACTTTGA | ,yAAAATATTT | TACCTCGYAG | 1980 |
| | ACGTCCCATA | TCATTaATTG | TAGCCGAATT | GTTGTTACCT | TTAATTTCCC | AAAAAGCTGG | 2040 |
| 55 | TACAGTAACC | TGATTAAAAA | ATCGTGGTTT | CATATTTTCT | GTATTATGAT | TATCTGCAAA | 2100 |
| - | | | | | | | |

| | TAAATCTTCT | TCCAACTTAC | TGGCTTTAAA | AGACTCATAT | AACTTTCGTG | AATGATCGTT | 2220 |
|-----|------------|------------|------------|------------|------------|------------|------|
| 5 | AAAGTAATCA | AATAATTTAA | TCATGTAGCA | CCTCTTGaAC | TAATGTTTCC | CATTTTAAAA | 2280 |
| • | TAATATCTTG | AGTCATAAAT | TGCTGTGCCA | CTTCATAAGA | GATGTCATGT | GGTGTCTGGG | 2340 |
| | GACCATTGTT | AAAATACATT | ACAATGGCAT | GAGCTAGTTT | TGCGATAACA | TCATCCACAC | 2400 |
| 10 | TATCTTCGTC | GGTATCAAAA | GGTACCAAGT | AGCCATTTTC | CCCATCTCGA | ATAAAGGTTG | 2460 |
| | GGTTACCATA | ATTCACATTT | AATCCAATCA | TACCTAGTCC | TGAGCCTACC | GCTTCCATTA | 2520 |
| | GTGTTAACCC | AAAACCTTCG | CTAGTTGATG | CAGAAAGAAA | TAACTCATAA | TCATTATAAA | 2580 |
| 15 | TTTCATCAAG | TTTAACATGC | CCTAGTAAAC | GAATATAATC | TTGTGCGCGG | TGTGTATCAA | 2640 |
| | TAATTTTACG | CAGTCGCGTC | TTCTCGCTAC | CTTCTCCATA | AATATCAAAT | GTTAATTCTG | 2700 |
| | GCACTTGTCG | TTTAGCCACG | ATAACCGCTT | TGACAAGCCA | ATCAATATGT | TTCTCATTCG | 2760 |
| 20 | CTAAACGAGA | TGCACTAATC | ATCGCATATG | GCTTTCTTGA | TAATGTTGGA | TATGATAATG | 2820 |
| | CATCAATGCT | TCCCACAGGA | ATAGTATAAA | CACGTGGGCG | ATAACCTTGA | TATTGCTCAA | 2880 |
| 25 | ATTGTCGACA | AACCATATGA | TTTTGAATAT | CTGTTGCTGT | AATAAAGAAA | TCAATGTATT | 2940 |
| | TAGCTTTTGA | AAATTGATAT | TCATAATAAT | TGTTCCATAG | TATATGCTGC | TCACTCATCA | 3000 |
| | TATTATTACT | ATAATGATCA | GCATGAATCA | CAACACCAAC | TTTACTATCA | CCTTTATGCT | 3060 |
| 30 | GCAAAACAGC | CTGACCAATA | TCAGAAGCGC | GGTCTAATAT | GACAATATCG | TCTCGGGTTA | 3120 |
| | AATTCAATCG | TTGTAAAAAG | TATGCAATAA | ATTCCGTTTT | GTTATACAAC | ACCGCATCTT | 3180 |
| | CAAACACATA | TATAGAGCTG | TCTCCATCAA | TATATTCGTT | ATAAGCGATG | GAACCATCTT | 3240 |
| 35 | GATTATAAAA | TTGTCGCATA | TATAATTTCG | CTTTATTATC | AGCTGGTGCA | TAATACTCAG | 3300 |
| | AAAATATGCG | CGTATAACTA | TAAAAATCTT | TACGTACTAA | CATACTATTA | ATTACAAATT | 3360 |
| 40 | CTGCACGATC | CACAATATCT | TTTTGTTCAT | TTTGCAGATA | ACATGTTACA | AATGATGATT | 3420 |
| | TCCCATTAAA | ATATAGGCGG | ACTATCTTAC | CATTTCTTTC | TCTAAAACTA | ATGTCATGAC | 3480 |
| | CAAGCTCACG | TTCAATGTCA | TCTAACGTGT | ACGTTGTTGG | TGCTAAAGAA | ATATCACTAA | 3540 |
| 45 | AATACTGATA | CAACCAAATA | ACTTCTTGAT | CTTTAAACCC | AATGTTTTGC | GTTAATGTCT | 3600 |
| | GTATGTTCTC | TGACTGTATA | AAATCTAAAA | ACACAAATTT | AGTGTCTTGA | TTTGTACGTC | 3660 |
| | TCAATAATTT | AGCACGGTAA | GCTTGTGCAT | ATTCAACACC | GCTACTCGCC | CAGCCTATAC | 3720 |
| 50 | CAAAGTTTAT | ATTATATATT | GTCATGCGCT | ACCCCTTTTC | ATTTATGGAA | AATGTATAAC | 3780 |
| | TGGCATACCC | TCTTTATCAA | ATGTAATCAT | GCTTTGACAA | ATATTTTTCA | CCATTCTTTT | 3840 |
| 55 | TTTGATATTT | CGTGTCATAA | CTTCAAATGA | ATCTAAGGCA | ACTCTATGGT | ATTCAAAAAT | 3900 |
| J.J | | | | | | | |

| | GACTTGTTCT | AACCAACATG | AATCAATTGC | TTTCAAAAAG | ACTITITGAA | CGAAAATATT | 4020 |
|----|------------|------------|------------|-------------|------------|------------|------|
| 5 | ATAATAATAT | GCACTTTGCA | TGTTTTTACG | ATTCAAAGCT | AATTGCTTTT | CAAATTGCTC | 4080 |
| | TAATAAAAAT | GTCACTACTG | CTTGCTTATC | TTTAAAATTA | ACACAAGCCA | CATCTTTATT | 4140 |
| | AAATTGGAAA | CTTAAATTTT | GATAAATATA | CTCGACAACA | CGCGATTTTG | TTAGCACCTT | 4200 |
| 10 | TTCCTCATTT | ACAAACATTT | CAAATACATC | TTTAGCTAAC | GCTTTAAAAT | CTTGATTCTC | 4260 |
| | AGCATCATCT | ATTTCTAAAA | CTCGATTGCG | TTCCTCGTAT | ACAAGATCTC | GCTGTATACT | 4320 |
| | AATGCTTTTT | TCAAATTCAT | TAGCCATTTC | ACGAGCTTTA | ACCCCTTGTT | CTTCCGAGAt | 4380 |
| 15 | aCGcTGCGCT | TTAACTACAA | TTTGCTTAAC | TTTGCGATTA | AACAAATTAC | TTTGCGATAA | 4440 |
| | TCGTTGTGCA | TCTAATGAAT | ATAATTGATT | ATTTTCCGCT | AAATTACTAT | CGCTCCATCG | 4500 |
| 20 | CTTAACTAAA | TAATCATCTA | GTGAAATATA | TATACAAGAT | GATCCCGGAT | CCCCTTGTCT | 4560 |
| 20 | ACCAGAACGA | CCACGTAATT | GCCTGTCTAC | ACGGCTATTT | TCCATATGTT | CATGAATAAT | 4620 |
| | AACAGCTAAT | CCACCTAATG | CTTCGACACC | TTCACCAAGT | TTAATATCTG | TGCCTCGACC | 468Ö |
| 25 | TGCCATACTA | GTCGCAACAG | TCATGGAACC | AATTTGCCCT | GCTTCAGCTA | TCATCTGCGC | 4740 |
| | TTCTTTTGCA | ACATTTTGCG | CAATGAGTAA | ATTATTAGGA | ATATCCATTT | GGAATAATAC | 4800 |
| | TTTCGAAAAG | TATTCAGCCG | CTTCAGCAGT | TCTCGTTATG | AGTAAAACCG | GTCGCCCCGT | 4860 |
| 30 | TTCATGAAGT | TCAACTATAT | CATGAATCAT | CGCGATGTTT | TTCTCATCAA | CTGAACGAAA | 4920 |
| | CACTTTATCT | GGTTCATCGA | TACGTTGAAT | CGCTTTATCA | GTTGGTACTT | GTACGACTAT | 4980 |
| | TTTTGAATAC | AAATCAAAGA | ACTCTGATTC | GCCTAATTTT | CCTGTAGCTG | TCATACCTGA | 5040 |
| 35 | AAATGATTCA | AAAAGTTTAA | ATAAATTCTG | GAAGGTAATT | GTTGCCATAA | CACTTTTATC | 5100 |
| | TGTTGAAACC | TCCATACCTT | CTTTCGCTTC | AATAGCTTGG | TGAAGTCCAG | CTTGCAACTT | 5160 |
| 40 | AGTTCCCGGT | AACATACGAC | CTGTAATACG | GTCAATTAAA | ACAATATCAC | CATTATATAC | 5220 |
| | AAAGTAATCG | ACATTAGATT | CAAACAAATA | TTGTGCGCGC | AGTGCTAAAT | TAATATTACG | 5280 |
| | CACTAGGACC | ATCGCTTGTT | CGCTATATAA | ATCTTCAACA | TTAAAGTATG | ATTGTGCCGC | 5340 |
| 45 | TTCAATACCT | TGATTTAACA | GCCATATTTC | TTTTTTTGGTC | TTCTTCATTT | TAAAATGCAC | 5400 |
| | GTCTTCAATC | AATGTATCTA | CAAACTCTTT | CACAATATGA | AATAGATTTG | ATTGTAATCT | 5460 |
| | TGGTGCACCC | GAAATAACTA | ATGGTGTTTG | AGCAGCATCT | AAAATGATTG | AATCCACTTC | 5520 |
| 50 | ATCAATAATA | CCGTAATTTA | ATTGTGGTAA | AAATTTCCCT | TCCGCACTAT | CAGCCAAATT | 5580 |
| | ATCAATTAAA | TAATCAAAAC | CGAGACGTCC | ATTAGTTGTA | TATATAATAT | CATGTTCATA | 5640 |
| 55 | TATATTACGT | TTTTCCCCTT | TTTGATACTC | ATAATCCACA | ATATCAACAA | AACCTAATGA | 5700 |
| | | | | | | | |

| | TAATCATTCG | TTGTAATTAA | ATATGTTCCT | TTTCCCGAAA | GAGCATTTAA | ATATAAAGGC | 5820 |
|------------|-------------------------|------------|--|----------------|-----------------|-----------------|------|
| | ATCGTTGCCG | TTAATGTTTT | ACCTTCGCCT | GTTTGCATCT | CCGCAATGTT | ACCTTCATGC | 5880 |
| <i>5</i> | AATACAATCG | CTCCGATTAA | CTGAACTTCT | TTAGGATACA | TACCTAATAC | TCTCCAGCTC | 5940 |
| | GCTTCACGTG | CCACTGCATA | AGCTTCAGGT | AACAATGTAT | CTAGTGTATC | AACTCCTGAT | 6000 |
| 10 | GCTAAACGTT | CTTTAAATTC | TATTGTCTTT | TGTTTTAACG | CATCATCAGA | ATATGATTTA | 6060 |
| | ACTTCATCGC | TCCATGTATT | GaTGsGTTcA | CTATTTTTCT | AATCGACTTT | AGTCTTAATT | 6120 |
| | CGTTTATCGT | AACATCTAGT | TTATGTTTCA | TTTACTTCCC | CACCATTCAG | TTTCGATACA | 6180 |
| 15 | TCTAAGTAAT | CTAAAAATCG | TACTGGATTC | ATTAAACGTG | ACATATAATT | TAGATGTTTG | 6240 |
| | TCTTGCTCTT | CTTTAAAATA | AACCTCGACA | TTTGTATCTT | TTAGTTCATG | ATTTCCTGGG | 6300 |
| | ACATGTTCTG | TAAGCCATCC | TTTTAAATCA | TCATCTTCAT | GGCTTGTACG | ATACACTTTG | 6360 |
| 20 | CAACCCAAAT | GCTGAGCGAC | ATAAGTTGCA | AAAACATTTG | ACTTTGACCC | ATAACTAATC | 6420 |
| | AAATTAATAG | CCTTTAGGGT | ATCTTGACTT | TGCAAAȚCAT | TCTTTAGTTG | CTTAATATTT | 6480 |
| | CCCTCGATAT | TGTCGTCCAT | CCAACGTTCA | ACGAGCCAAA | CATGACCAAA | CAGTTTCAAA | 6540 |
| ?5 | AAATCATTCG | AAATAGTTGG | ATAGGTGTCA | GATGGTTCTG | CAATAATGAC | ATTGATCATA | 6600 |
| | TCATTTCCAT | ATTGGTCATC | GCCTATCTTC | GTCACCCGCA | TGCTTTTATA | CTCTAAATCA | 6660 |
| 3 <i>0</i> | TATTGATGCG | TCATCTCTGT | GATTGTTAAA | CATCTAAATA | TAAGACTCGT | CGATGCTGCA | 6720 |
| | TTCATCATTT | TTATTTTATA | AGCATAGGCk | TCATCAGGAT | ATTGAATCGT | AATACTATTT | 6780 |
| | GACTTTACAA | TCTCAGTACT | TAGTTTTGTG | CCATTTTTAT | TATAAAAAAT | GATGATAAAA | 6840 |
| 35 | TACACTGAAC | CAGCAGGCGT | TGCATCAAAA | TCAAAATGCA | ATTTATAATG | CTGTCCTCTA | 6900 |
| | CGCAAAATTG | Gkaaacttgg | CGCACTTTTA | TATTTTGAAA | ATTGCTTTAA | CATCAACCAC | 6960 |
| | TCATGAATCG | GTAATCCAGA | GGGCATCAAA | GGATTTATAA | AAGTCACTTC | ACCATTTGAA | 7020 |
| 10 | AATGATACTT | TAGAGCCATA | CATAAATGTA | GTTTGTGAAA | TATAATTCCA | AGTAACTTTA | 7080 |
| | AATGTTTTGT | TTTTCAGCAT | GTTGAACTCT | CCCAAACTTG | TCTTCCAAAA | TAATGTTGTA | 7140 |
| _ | AAAATTAACA | AACCAACTTG | CAATGGTAGG | TGAATCATCA | TTATGTCGCC | CAGGAATACT | 7200 |
| 15 | GCGATTCATC | ACTCTTGCTT | GGTGTGCTGT | CAATACAGGT | AATAGCTCTT | GAAATGCATG | 7260 |
| | TGGATCATAA | TCATCATGTT | GCATATATGC | TATGGCAAAA | ACAGTTTGTG | ACAATGATTY | 7320 |
| 50 | CTTTTGAAAT | GTTTGCCAAA | ATTTTTGATT | TAATGCCTGT | ATCGACGCTT | GAGATGTATC | 7380 |
| | ACCTTCATTA | GACACCAGGA | CGTCTAATGC | TGTACCGAAC | TCTTCTGGTC | TAAGTAATCG | 7440 |
| | <i>ር እ ጥ አ ጥር</i> ምምር እ | CCAATCCTTC | ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ | N ACHTOCHETERN | CC3 5 C3 5 T3 5 | TTTC CCTTC N CC | 2501 |

| | TAATTCATGT | GATTTAAAAT | TCAGCTTTTC | TAATGTCTCG | TCAATAACAT | TGATAATACC | 7620 |
|------------|------------|------------|------------|------------|------------|--|------|
| | TTGTTCATAT | TCAGATGAAC | CGATATAAAA | ACTACCACCT | TCAACACGAG | GATCGCCGAT | 7680 |
| 5 | AAGTAAAAAC | GGTGCATTCA | TACGTTTCAT | CATATAATAT | CCTTCGAAAC | CTTCCGCTGT | 7740 |
| | TCGATAACCA | СТААААТАТА | CGTTTAGTGG | CGGTTTCATA | TCACCAGGGT | GGAAATAATA | 7800 |
| 10 | AATAAATTCC | TGTCGTTGAC | TATCTACGAA | ACGACTACCA | CCAAGTAAAA | ATTGACCCAT | 7860 |
| | GTCTAATCTA | GACCATCGTT | TGTGTATAGG | TCCTAAATGT | ACCGTCCCGT | TCCCACGCGC | 7920 |
| | CTTAACAGTT | ACACTTATAT | AAGCATCAAA | TGGTTTCGCA | GGTATCTCTA | AAGGACTGTC | 7980 |
| 15 | TAACATATCA | TCAGTCAATA | CGATTTGTTC | AATTAATGCA | CCATCAGCGC | CAGTCTGAAT | 8040 |
| | CAATCTAAAT | GTATATTGCA | ACTCGACCGC | ACCATCAATA | TCAAATTCTG | GCCATATTTG | 8100 |
| | AATGACTTTA | TCTTTATCGT | AAACGAGATT | ATTTTGCCAA | GATGCGATAG | GTTTAAATTC | 8160 |
| 20 | TTTCCCAAAT | TCTCCACTCA | ATGTGAGCTC | TGAATTACCT | TGGTAAACGA | CATCTCCTTT | 8220 |
| | AAAATTCGGA | TGCACAAGTG | CTAACTTAGG | AGAAACCTTA | TCTCCATACT | GTCCTGAGAA | 8280 |
| | GCTAACTGCC | TCTAATTTAT | TATTACGTTC | TTCAATATTC | CGGTAATGTA | ATGGTTGAAC | 8340 |
| 25 | AACGTATTTT | TGGACATTTT | CGTCTTGTTC | ATATTCAACT | GACCAAAATG | ATTCATCAAC | 8400 |
| | ATACGTATTG | TATGGTTCGC | TTATCATTTG | TAATAAATTC | GTTAATGTCT | CCGAGTATGG | 8460 |
| 3 <i>0</i> | TGCTTGAATA | TAGATAAAAT | CAAAGCGCCC | TTCTGCTTCA | ACAATCGCTT | CAATAGCCTC | 8520 |
| - | TACATAACCA | CTATCAAATT | CAAACAATCC | AATATCGAAG | TAATCCCAAC | TCACACCTTT | 8580 |
| | TTTGTGTTGA | AAAATAGGTT | CTAAATCGTC | TCCTCCAATT | TGCAAAACTC | TAAATTTACG | 8640 |
| 35 | TGGCATCATT | TTCACCTTCT | ATTAACTCAT | CGAGCTGATT | AATAATATTC | TTAGAAGCAT | 8700 |
| | ATGCATCTAT | TAATTTTAAA | GAATAGGCGT | ACGCATAATT | CCAATTTTTC | AAATAAAATA | 8760 |
| | TTAATAĀTAA | TAACGCATCA | TCTAATTCAT | CAACTGTATT | TATAATACGG | CCATTGTCAT | 8820 |
| 40 | AATCAGAGAC | GTAATCTGTT | TGTTGACCAT | TAATTTGTGG | AATCCCAGCG | CTAATTGCAC | 888 |
| | TAATTTGTAA | ATACAAGTCA | GGTTCTTTTG | ACATATCTAT | CACAAGTCGC | AACGTCCGCA | 8940 |
| | ATGCTTCTAC | AACATCATGT | TCAGCATGTA | TCGTCTTAAC | AGCAATGATG | TCATCTTGAT | 9000 |
| 45 | CTTCAGGTGT | CATTAATGCT | GAAACATTAA | CATCCGCATT | CTGTTTAGCT | TGGTATTCCT | 9060 |
| | CATTTACCGA | CGTAATACAT | TCACGAAGCC | ACATCGGTAT | GTCATTTTGA | TGGCGCGATA | 9120 |
| 50 | ATAAAATTAA | ACGGTAATAA | TCTTCCTGTG | CGATATAATC | CACAAGTCGT | TGCATCATTT | 9180 |
| | GTTGCAAATC | AGCGTCACTC | ATACCATCTA | TCCATACACC | TATAAATGTT | TCCATCAATT | 9240 |
| | САСТАСТТАТ | ATTACCTCAT | тетстетт | CAAATGGTGT | CATTCGAATC | ערייייי איייייע איייייערע אייייי אייייי | 9300 |

| | TTAAATGGGC | ATTCTTTACG | ATAGATTGAT | ATTCCTCATC | TGACACAGTT | TCATTTCTAT | 9420 |
|----|------------|------------|------------|------------|------------|------------|--------|
| | TTTTAAAAAA | TGAATAACTT | AATGATTTCG | CTGGAATATG | ATTGGCTATT | TGTCGATTGT | 9480 |
| 5 | GCCTAGCATC | TGAAGCCACA | ATCACATGAT | CATCTTCATG | TATTTGTTGT | GCAATCATTG | 9540 |
| | CTTGAAATTT | TTCTTCAATT | AGTTGAGCCA | TATTGTTATA | TTCTGTTTGT | TGATAGTGAT | . 9600 |
| | GTTGATATCT | TTTTGAAACA | GTGACTCTGC | CATTTTTCAA | ATCTTCATGA | AGTACACAAT | 9660 |
| 10 | CTCCATTAAT | CGTTAAATAT | TCTTGGTAAG | AAGCCTCTCC | CTGATCATCA | AAATAACGTA | 9720 |
| | TCGCTGATAA | ATAACCTCTG | TCATCAAAAA | TATAACGCCG | TTGTAACTGA | TCTCTTTCAA | 9780 |
| 15 | ATTCTTCAAA | CCAAATTGAA | TACCCTTCTT | GACTAAAATA | AATATTTGTA | TAGGTCTGTT | 9840 |
| • | CACTCGTCAC | ACATTTTAAT | AAATACGGTG | TGTACACAAA | CTCAACATCA | TCCGGCCATT | 9900 |
| | TTAAGTGATG | ATAATTAATC | GCTTGTGGCG | CATGGTGACT | GAATCCTTGA | ATTTCATCAA | 9960 |
| 20 | ACACAGACGA | ATACTTTGTC | TCATATAAGT | CATATCGATG | TAAAAATGTT | CTTAAATTTG | 10020 |
| | GTGCATGATT | GAGAACAATC | AGTTGATAAT | CTAAGTCATT | TTCAAGGTGC | ATTCCCATTA | 10080 |
| | AACTAATCAT | ATCGTCAAAT | TCCGTCTTAT | TTTGTAGTTG | ATAATACGGC | ACAGTCGTGT | 10140 |
| ?5 | CTTGCCACCA | TCGTTGGTCA | TCGTACCAAG | CTGGAATAAA | GTATTTCATA | ATTACCTCCT | 10200 |
| | TACCAATACT | GGTTTAAAAA | TGGCTTATAT | TTATCAAAAT | ATAAATATGT | ACGAATTGTT | 10260 |
| 30 | TCTGCAATAT | TAATACTGAT | GTAAACTAAT | ACAATCAGTT | GTACTGAGAA | ATAAATTTCA | 10320 |
| | GTAGATAAAT | GCGGTACAAA | CAATGTGAAA | TAAAGCGGTA | TACCAATAAT | GACTGTAACT | 10380 |
| | AATGCCAATC | CAAACCAACA | TACGCGTCGT | GCTTGATAAT | TTAAATAACG | TTCTGTATCC | 10440 |
| 35 | TTACCAGGTT | TaACTCCTGA | AAAATAATTG | CCACTCTTTA | AGAAATCTTT | GGATTTTTGT | 10500 |
| | TTAGTATTGA | TTAAAAATCT | CGATAAAAAA | TAACCCAATA | ACATTTGAAT | CACTAAATAT | 10560 |
| | actgāaatac | CTACTGGACT | ATCAAATGTC | AGCATTGGCA | TGTCATCTGA | TATGCTTTTA | 10620 |
| 10 | TTAAACATAG | ATAAAATAA | ATGAATGCCA | CTTTTTAAGA | AAACAAAAGC | TGAAATACTC | 10680 |
| | ATCATTAAAG | TAATACTGCC | TGCAGGGTTA | ACTITCCAAG | ATAAATAAGA | TTTCATATTT | 10740 |
| | GTTGCGGAAA | CGTTCATTAA | ATCGATATAT | GGTATTCTCA | CTTCTACTAA | TTCAATAAAT | 10800 |
| 15 | AATAAGATAA | ACAATGTGAT | TATCACAAGG | ATGATTAACA | ACGCAATCAC | AATATGACTT | 10860 |
| | GCATCTATAT | ATTCCATTTT | TTGATGCATC | ATTGATTTAA | TAATACTAAC | CATTACAATC | 10920 |
| 50 | GGCATTGGTC | CTGCGATGCC | GTAGCGACTA | TTTTTGTCAG | CTAACCAAAC | TAATAACATC | 1098 |
| | GTTCCAGTAA | CCAAAATCAA | TATTGTTAAG | TAAATATTGT | CTTGATGAAC | ACGTTCTTTC | 11040 |
| | GAAACATATT | CATGAATCAC | AAAATAACTT | TGAATAACAC | TTAAAATTAA | TGTTAAGATG | 11100 |

| | GAAATCAGCA | TCAAGATAAT | CATTGATGTT | AACCACGGAC | CTAALCCTAA | AGTGAAAATG | 11220 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | TTTAAAGTAT | TAACGTCTCC | ACCCATATTA | GAAATAGCTA | TTTTAAAAAA | TGACTCATGT | 11280 |
| 5 | TTTACTTGCA | TATCGTTaTA | GGAAACGATG | GAAATGTTTG | TGCCTAATAT | ATAAATAaAC | 11340 |
| | AAGATAAAAC | ATGTGTATAG | CATACGTTTA | TATATAATTT | TATATTCGTA | TTGTTGTAAA | 11400 |
| | AGTTTTAACA | TGTTGCACCT | CTTTTATATC | AAAAACATTA | AAAAGACTAA | GGGTTCATCA | 11460 |
| 10 | CTAATTATTA | AAATCCTATA | TCGATTTTTC | TAGTGATTGG | TGCCTCAGTC | TTTTTAATTT | 11520 |
| | TAGCCAGCTA | TAAATTCAAT | TTATGCTTGA | GAATCATCTT | GATCATTTTC | ATCTTTCTTT | 11580 |
| 15 | TTCTTTCTCT | TCATTAAACC | TAAACCAACT | AATAATGTCA | TAACGCCACC | TAGTAATCCA | 11640 |
| | TTTTGTTTTA | TTGAGTCACC | TGTATCTGGC | AATCTTTTTT | CACTTTGTGC | TGGTGTGCCA | 11700 |
| | TTATGTTTAG | TCACTTCAGA | TGTTGCACTT | AATGTAGACT | GAGATTCACT | CGTGCTCGTT | 11760 |
| 20 | GTTGCTTCAC | TTGATAAGCG | AGATGTGCTC | GTGCTGTGAG | TATGATGCAT | ACTCATTGAG | 11820 |
| | TCTGACGGAT | GCATTGAGTT | AGATTCAGAT | GTACTTGTTG | AGCCGGACAT | ACTTGTTGAT | 11880 |
| | GTTGAGTCAG | AAATGCTTTG | TGAACCAGAC | ATAGATGTAC | TCAGTGATTC | GGATGTGCTT | 11940 |
| 25 | GTCGAATCGG | ATGTGCTCAA | TGACGTTGAT | GTGCTTGTTG | ACACTGATTC | TGAGTCACTA | 12000 |
| | ATTGATGTTG | AGTCGGATTT | GTCTTGTGAC | ATTGAAACAC | TCGATGAATT | AGATTCACTC | 12060 |
| 30 | ATTGATGTTG | AGTCAGATAC | GCTCGTTGAA | CCTGAACCAG | ACGTACTTAA | TGATTCAGAT | 12120 |
| 30 | ATGCTTGTTG | AAGTTGAACC | ACTTGTTGAG | TCCGATGTAC | TTGTCGATGT | CGAGTCTGAA | 12180 |
| | TCTGATGTAC | TCAATGATTC | TGAGTCACTG | ATAGAAGTTG | AATCACTTGT | AGATTCTGAT | 12240 |
| 35 | TCTACTGTAC | TTTGTGAACC | ACTGATACTT | ATTGAAGTAG | AATCACTGAT | ACTGTCTGAT | 12300 |
| | GTTGATAATG | ATGTCGACAC | CGATGTGCTT | TGTGATGACG | ATGTACTAGC | ACTCATTGAC | 12360 |
| | ATTGATGTTG | ATATCGATGT | ACTTAAGGAA | CCAGATGCAC | TTGTACTTGT | TGACTGGCTT | 12420 |
| 40 | TGTGACATTG | AATCACTTAA | TGATGTAGAT | GTGCTTGTTG | AGCTCGAGTC | ACTTACACTT | 12480 |
| | GTTGAACCTG | ATATTGAGTC | ACTTAAACTT | GTCGATGTTG | AAACTGAtwC | GcTTCCGCTC | 12540 |
| | ATTGAGTCAG | ATGTTGAAAG | TGATGTACTC | GTTGAATTTG | ATCCACTGAT | GCTAGACGAA | 12600 |
| 45 | TCACTTGTAG | ACATTGAGTC | GCTTTCTGAT | GCACTGATGC | TCATAGAGTC | AAATTGACTA | 12660 |
| | TTACTTGTTG | AGCTTGACTG | CGAATCGCTC | ACACTTGTTG | ACGTTGATTC | TGATCCACTC | 12720 |
| 50 | ATACTTTGCG | AGCTACTCAA | TGATTTTGAA | TCACTTAATG | AATCCGAAGT | GCTAAGACTT | 12780 |
| | GTGGAACCAC | TTAAAGATAT | TGATCCACTT | AATGAGTCGG | AGTCACTTGT | ACTAGTAGAA | 1284 |
| | TCACTCATTG | ATATTGAATC | ACTTAGCGAG | GTAGACTYGC | tTACGCTTTC | TGAACCACTT | 1290 |

| | TITGATICAE TIATGATIC AGAITCACTC ACGCTTTCTG AACTTCTTAG TGACGTCGAT | 13020 |
|----|---|--------------|
| | ACACTTAATG ATGACGAATC GCTTGTGCTT ACTGAATCG | 13059 |
| 5 | (2) INFORMATION FOR SEQ ID NO: 221: | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 10758 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| | | |
| 15 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 221: | |
| | AGGGATGGCC TTACCTAAAA AACCGGGNAA ACCCTCCAAA ACCCATTAAA AGGNTGGNTA | 60 |
| | CCCTTTAAAA TGGTAGCATT TAACCGCCAC CCGCCAAGGT GGGTGGTTTA TTCTTCCGTT | 120 |
| 20 | ATTTAAATTA GTACACCATG CAGATTCTGT AGTTGAGGGA TATTTTAACG AAAGCTTATT | 180 |
| | AGCAACTGAT AAAAAAATAC GTCCTAAGGC ATATATTGCT TCATGGAAGG ACATCGAGCC | 240 |
| | GGCTAAGAAA ATAGAATTTA AAATTAAAAA AGGTATTAAA TGGCATGATG GTAATGAATT | 300 |
| 25 | GAAAATTGAT GATTGGATTT ATTCAATTGA AGTCTTAGCT AACAAGGACT ACGAAGGTGC | 360 |
| | TTATTATCCA AGTGTAGAAA ATATCCAAGG TGCGAAAGAT TATCATGAAG GAAAAACTGA | 420 |
| 30 | TCATATTAGC GGATTGAAGA AAATAGATGA CTACACTATG CAGGTTACAT TTGATAAAAA | 480 |
| 30 | ACAAGAAAAT TACTTAACAG GATTTATTAC TGGACCTTTA TTAAGTAAAA AATATTTATC | 540 |
| | AGATGTACCA ATTAAAGATT TAGCGAAATC AGATAAAATC CGAAAATATC CTATTGGTAT | 6 0 0 |
| 35 | TGGACCGTAT AAAGTTAAGA AAATCGTTCC AGGTGAGGCT GTTCAACTCG TTAAATTTGA | 660 |
| | TGATTATTGG CAAGGTAAGC CTGCACTAGA CAAAATCAAT TTAAAAGTTA TTGATCAAGC | 720 |
| | GCAAATTATT AAGGCAATGG AAAAAGGCGA TATTGATGTT GCGAATGATG CTACCGGTGC | 780 |
| 40 | AATGGCAAAA GATGCTAAGT CATCTAATGC TGGTCTCAAG GTATTATCTG CGCCAAGCTT | 840 |
| • | AGACTACGGT TTAATAGGET CGTATCTCAT GATTACGATA AAAAAGCTAA TAAAACTGGT | 900 |
| | AAAGTGAGAC CAAAATATGA AGACAAAGAA TTACGTAAAG CAATGCTTTA TGCAATTGAT | 960 |
| 45 | AGAGAAAAT GGATCAAAGC GTTTTTCAAT GGTTACGCTA GTGAAATCAA TAGTTTTGTA | 1020 |
| | CCATCTATGC ATTGGATAGC AGCCAATCCT AAGGACCTAA ATGATTACAA ATATGATCCT | 1080 |
| | GAAAAAGCTA AAAAAATCTT AGATAAGTTA GGTTATAAAG ATAGAGATGG TGACGGATTT | 1140 |
| 50 | AGAGAAGATC CTAAAGGTAA TAAATTTGAG ATTAACTTTA AACATAATTC AGGTTCTAAT | 1200 |

55

CCTACTTTTG AACCAAGAAC TGCTGCGATA AAAGATTTCT GGGAAAAAGT TGGCTTGAAA

| | AATACGATTC | CTGTTTATAT | GCCATATATC | ACATCTTATT | TCATGACGCG | TGCTATCGGC | 1380 |
|------------|------------|------------|------------|------------|-------------------|------------|------|
| | GACAGACCTT | TAGTCGTCCC | GCATCAATCT | CAGAACTTAG | CATTTATTGG | TAACTTTGCA | 1440 |
| 5 | GAAACAGAGC | GAGACACTGT | ATTTACAACA | GAATATTCGG | TTCGTACTGC | CATGGAAGCT | 1500 |
| | GTTTATCAAT | TACTAAATAT | AGATCGTGGT | ATTCCAGAAG | TCATCAATAG | TCCATTTGAT | 1560 |
| | CTTCGCGTCT | TAATGGATGC | CATATACGAA | CTGAATGACC | ACCAAGATTT | GCGTGAGATT | 1620 |
| 0 | ACTAAAGATT | CGAAAATGCA | AAAACTCGCA | TTAGCAGGAT | TCCTTAAAAA | GATAAAAGGT | 1680 |
| | ACGTACATTG | AGTCATTATT | AAAAGAACAC | AAATTGTTAT | AACGAAAACC | ATTAATAGAT | 1740 |
| 5 | TTTTATTTGG | TGATTTCAAA | TCATGAGACT | GGGACAGAAA | TGATGTTTTC | АТАААААТТА | 1800 |
| | TTTCGTTGTT | CCACTCTCAT | GATTTTTTTG | ATGAAACATA | ATTACATGAT | TGATTGCATC | 1860 |
| | ATTTTGTTAA | ACAAGTGATT | GCAAACCTGC | CATTTCACAC | TGAAAATTTA | CATAATAAGT | 1920 |
| 20 | GACGATATTT | TACAAGTCAT | ATACAAATAA | CATATATTGT | TAAATAATTT | TACCTAATCT | 1980 |
| | TAACATTAAA | TTTACAATTA | TAAGCGATAA | TCTAAATATA | AAGCTTATTT | GAGGTGAAAT | 2040 |
| | AATGGAAATG | TCGGTTACAG | AAGTCATTTT | CTCCTTTTTA | GGTGGTTTAG | GTATTTTCCT | 2100 |
| ?5 | TTACGGCTTA | AAAATCATGG | GAGACGGGCT | TCAAGCATCA | GCAGGAGACA | GGCTACGAGA | 2160 |
| | TATTTTAAAC | AAATTTACAT | CAAATCCAGT | ATTAGGTGTT | ATTGCAGGTA | TCGTTGTAAC | 2220 |
| | TATTTTAATA | CAAAGTAGTT | CAGGTACGAC | AGTTATCACA | ATCGGACTGG | TAACAGCTGG | 2280 |
| 30 | ATTTATGACA | TTGAAACAAG | CCATTGGAGT | GATAATGGGT | GCTAATATCG | GAACAACGGT | 2340 |
| | AACTGCATTT | ATTATCGGTA | TAGATTTAGG | CGAATATGCA | ATGCCAATTT | TAGCATTAGG | 2400 |
| 35 | TGCATTCTTA | ATCTTTTTCT | TTAAACGCTC | TAAAATCAAT | AACATTGGCC | GCATACTATT | 2460 |
| | CGGTTTCGGT | TCACTATTCT | TCGGTCTAGA | ATTTATGGGT | GATGCCGTTA | AACCTTTAGC | 2520 |
| | ATCATTAGAT | GGATTTAAGC | AATTAATGCT | TGATATGTCT | ACAAATCCAA | TACTCGCTGT | 2580 |
| ‡ 0 | CATTGTCGGC | GCAGGGTTAA | CAGCACTAGT | TCAAAGTTCA | AGTGCGACGA | TTGGTATTTT | 2640 |
| | ACAAGAATTT | TATCAACAAG | ATTTAATTAG | CTTAAACGCA | GCAATCCCTG | TGTTACTAGG | 2700 |
| | CGATAACATT | GGTACCACGA | TTACAGCTAT | CTTAGCTAGT | TTAGCCGGCT | CAATCGCTGC | 2760 |
| 45 | AAAACGTGCG | GCGCTTGTAC | ACGTCATCTT | TAACTTAATC | GGGGTAATTA | TCTTCACAAT | 2820 |
| | TTTCTTGCCA | GTTGTGATTC | ATTTGATTAG | TTTGTTACAA | GATTTATGGC | ACTTAAAACC | 2880 |
| 50 | AGCGATGACG | ATTGCAGTAT | CACATGGTAT | CTTCAACATA | ACAAATACTT | TGATTCAATT | 2940 |
| | ACCATTTGTA | GCAGGTTTAG | CATGGATTGT | TACAAAGCTT | GTCCCAGGTA | AAGATATTGC | 3000 |
| | TGATGACTAT | AAACCTCAGC | ACTTAAACAA | AGATCTTGTT | TATCACGCAC | CTGGTGTTGC | 3060 |

| | AGACATTCG | GAAATTACAA | AAGACGATA | AAAATTGATC | : AAAAAGCTTG | AACAAAAGCA | 3180 |
|----|------------|--------------|------------|--------------|--------------|------------|------|
| | TCAAGCTGTT | r gaaacaatca | ATGATAGCAT | TCGAAATTAT | TTAGTTAGAA | TTTCTACAAA | 3240 |
| 5 | AGCCATTACO | AAGGCAGACG | TTGAGCGTTT | AGCAGTTATG | TTTGATGTCA | ATCGCTCTAT | 3300 |
| | TTTAAAAGTA | GCAGAGCTAA | CAGAAGAGTA | TGTCGCTCAA | TTAAAACGCC | AACATGATGA | 3360 |
| 10 | AGATATTCGC | ATTACAGAAG | ATGCACAACG | CGGTATGGAT | AAATTATTCA | ACCATGTTGC | 3420 |
| | TGAGTCATTT | GATAAAGCCA | TCGACATGTT | ' AGATGTTTAT | GACAAAACGA | AAAAAGATGA | 3480 |
| | AATTGTAGAA | CGTAGTAGAG | AATCATTTAA | TATTGAACAT | AAACTACGCA | AAGGTCATAT | 3540 |
| 15 | TAAACGCCTT | * AATCGTGGTG | AATGTACAAC | AAAAGGCGGA | TTACTATATA | TCGATATGAT | 3600 |
| | TGGTGTTCTT | GAACGTATCG | GTTATCATTC | ACGAAATGTT | TCTGAAGCAC | TTGTTGGCCT | 3660 |
| | TAACGATGAT | GTACCTACAG | ATGAAGAAAT | TGCAACAACT | GAAATTTAAT | TTTTACTGTC | 3720 |
| 20 | TTATTTATAT | TCATATTTTT | TTAAAATTAG | AGATTCAGAT | GCATGTAAAA | AGCCAATCCA | 3780 |
| | ACATTCATGG | GTTGGCTTTT | TTGTTTAGCA | AAATTTATTA | TCTTAAATCG | GCTATAAACA | 3840 |
| | CTGATATAAT | AATGCTTCAT | TAGTATGCGG | TAAGCATGAC | GGACACTGTT | CTCGGAGTCT | 3900 |
| 25 | GACCCCGAAA | CGTTTAATAT | ACACTTTTAC | ACGTCGCCTT | CATTGAAGCG | AATTGCCATA | 3960 |
| | ACCTTCACAT | TATATATAGT | TCTTTCCATA | TAAATGTCCA | AATTTTTAGA | ACAACGCAAT | 4020 |
| 30 | AAATAACCAT | CCACCTAACT | TATCAAAAAT | TTAAGTGGAT | GGTTTTTCAT | TTTCATTTAT | 4080 |
| • | ATTTATATTA | GTGTTAATCC | AATCATAGAT | TTATCTATAT | GCACTGCTCT | ATACATTTCC | 4140 |
| | TCATTTAATT | TGCTTTACTT | TCATTTATAT | CATTATCAAA | ACACTTGGCG | TGTCATCGTT | 4200 |
| 35 | ATTATTTCGC | ATCTTTGACA | CGTTTATCAT | CATTAGGAAT | CGCGAATAAA | ATTGCGATAA | 4260 |
| | ATGCCATGAT | TCCCATTAAT | ACGTTAACCC | AAAGTGCAAT | CATCGCACCT | GTATGAATGC | 4320 |
| | TCGTTGCAGC | AACTGCACCA | GCATATACAG | CACCACTAAT | TGCGACACCG | AATGCGCCAC | 4380 |
| 40 | CAAGTGATGA | AGCCATTTTA | TAAATACCTG | AAGCAACGCC | AACTTTATCT | AACGGTGCAT | 4440 |
| | TCGAAATAGC | TGTATCTGTA | GAAGGTGTTG | CATAAATACC | TAAGCCTAGT | CCGAAACATA | 4500 |
| 45 | AATATCCTAC | GACACAACTG | ATAACATAAA | ATATGCCTGG | TAAGAATACT | AATGAAATAA | 4560 |
| 45 | GTGCAATACC | AATGACCACA . | ATGAATGTAC | CTAATAACAT | TGGTCGCTTA | GAACCCATTT | 4620 |
| | TTTGTAATAA | TTTTTCACCA . | ACTCGAATCA | TCAATAACAC | CATGATTAAA | TAAGTAATTG | 4680 |
| 50 | ATAAGTATCC | TGCCTGCAAT | GCTGTATAAC | CTAAACCTTG | TTGCACGAAT | GTATTCGCTA | 4740 |
| | CAATTAATGT | ACCTGCAAAA | CCGTTTAATA | AGAAGTTCGA | AATCGTTGCA | CCTGTATATG | 4800 |
| | GTTTATTTTC | AAATAATTTA | AAATCAATAA | GTGGATTATC | TACTITITIC | TCAACATTTA | 4860 |

| | AACCAAGTGC | TGCACCTTTA | GTAATGACAA | CGTTTAAACT | TAGCAACATA | ACTACTAGAA | 4980 |
|------------|------------|------------|------------|------------|------------|------------|------|
| | CAATTAGCCC | TGCAACGTCA | AATTTATGTG | TATTGGTAAT | TTCTGATTTC | GTTTCAGGCG | 5040 |
| 5 | TCCCTTTGAT | GAGTAACATT | GAAAGTACGG | CAACGATAAT | TGAGAAGATG | AAAATCCATC | 5100 |
| | TCCAACCCAT | AGTTGTCGCA | ACTGCACCAC | CGAAGAGTGA | ACAGATACCA | CTGCCACCCC | 5160 |
| 10 | AAGAACCGAT | AGACCAATAA | CTTAAGGCAC | GCTGACGTTC | AGCACCCTGA | TAATAAGTTT | 5220 |
| | TCATAATGGC | CAATGTAGAA | GGCATAATAC | ACGCTGCTGA | TACACCTTGT | ATAACACGAC | 5280 |
| | СТАВАВТТАВ | TAATGCCGGT | AAATTCGTAA | TAATAATTAA | TGCTGAACCA | ATAATACTTA | 5340 |
| 15 | ATAATAAACC | GATATTCGTC | ATTTTCACGC | GCCCAATTTT | ATCTGCCAGA | CCACCTGCTC | 5400 |
| | CAACAACAAA | CATGCCTGAA | AATAGTGCAG | TTAGACTGAC | CGCAATACTA | ATTGTCCCCA | 5460 |
| | TGTCTGTACC | AAAACTTTGT | TGTAAATTCG | GTACAACATT | TACAAGTGAT | TGTGCAAACA | 5520 |
| 20 | ACCAAAATGT | AATAACACCT | AATACAATAC | CTAAGATTAA | CTTGTTGCCC | CCGCGATACG | 5580 |
| | TTTCATTCAT | GTTAGTTATC | TCCTTTAAGG | ТААТСТАААА | CAACTGTCCC | TACTGCTTCT | 5640 |
| | GCAGAAATAA | GTAATGATTT | TTCTGAAATG | TTAAATTTAG | GATGATGATG | TGGGTAAATT | 5700 |
| ?5 | TCACCATTTT | CCACCGCTGC | ACCTGTATAA | ATAAAGGCAC | TTGGGCGTTC | TTTAGCATAA | 5760 |
| | TATGCAAAGT | CTTCTGAAGG | TGGTTGTGGT | TCACACATTT | CAACACCAAA | ATCAAGGTTT | 5820 |
| 3 <i>0</i> | GCTTCTTTCA | ACGTCTTAGC | CACGTACTCA | GTAAACTCTG | GATCATTATA | TAATGCTGGA | 5880 |
| | TAATCATCGT | TATATTCTAA | GGTGCAAGTt | ACACCATACA | TATCCTCTAA | TCCTTTTGAT | 5940 |
| | AAACGTTTAA | TTTCTTTTTC | AATTGTTGCT | TTTGTAGCAT | CTGTTAATCC | ACGTACATCA | 6000 |
| 35 | CCTTCAATTT | CAACAACATC | TTTAATGACA | TTGAATTGAC | CTTTACCGTC | AAATGAACCG | 6060 |
| | ATTGTGACAA | CACCGGTTTC | AAATGGACTT | AGTCGTCTAG | ATACAACTGT | TTGTAACGCT | 6120 |
| | GTGACGAAGT | AGCTACCTGC | AACAATGGCA | TCATTGGCCA | TATGTGGTGA | TGAACCATGA | 6180 |
| 10 | CCACCTTTAC | CTTGAACTTT | CAATTTGAAG | AATGCGCGTC | CTGTTTGAAC | ATAACCAGGT | 6240 |
| | CTGTAATACA | CTTTACCTGT | TTTCATTGTG | CTCATGACGT | GTACACCTAA | TACATGATCA | 6300 |
| | ACACCGTCTA | ATACACCATT | TTCAATCATT | GTTTTAGCAC | CACCTGGTGG | TACTTCTTCA | 6360 |
| 15 | GCTGGTTGAT | GTATCACAAC | GACTTTTCCT | GTAAAACTAT | CTTTCATTTC | AGCAAGCGTC | 6420 |
| | TCTGCTAATA | CAAGCATGTA | TGCTGTATGT | GCATCGTGAC | CACATGCGTG | CATAACACCT | 6480 |
| 5 <i>0</i> | TTATTTTGTG | ATGCAAAAGA | TAATCCTGTA | TCTTCAGTAA | TGGGTAATGC | GTCAAAGTCT | 6540 |
| | GCACGGATTG | CTAATGTTTT | ACCAGGTTTC | CCTGAATCAA | TCGTTACTTT | AATTCCACGT | 6600 |
| | GGTCCGACAT | TCGTTTCTAC | TTCCACATCT | TTACCTTTGT | AAAATTCAGC | GATGTATTTC | 6660 |

983

| | ATCATTTTGC | CTTCTTTAGA | TTTTAAAGTT | TCAATTAATT | GTTGATTCAT | ATCCTTCATC | 6780 |
|----|------------|------------|------------|------------|------------|------------|------|
| | TCCTTAGTTA | CATCATAAAT | GATTAATCAT | TATTTATATT | GCCAACAACA | GAGATGTTAA | 6840 |
| 5 | CCATTAATTT | TTTGCAATTT | TAGCTTTGAA | TATAAAAAAT | CACAAATTAT | GTATATCAAA | 6900 |
| | ATTTGTGATT | TGTGATCATT | TTATGAACTT | GGGTAACGTT | TTACTTCAAT | TAAGTGAATC | 6960 |
| 10 | CCATTCGTAA | TCATTTTAAT | GTTTAATGCC | AGTGTGTCCG | TGATATCTAT | ATCATATACT | 7020 |
| | TCTAATTTCG | GAAAACTCAT | TCGATTAACG | TAATCTATAG | AGTCCTTGTC | CATGCCATGT | 7080 |
| | ATCGTATGAT | GTTTGCGCCA | AAGATTAAAT | AACGCACCAT | TTTCTTTATC | TAAGGTAAAA | 7140 |
| 15 | TGTTTAATCT | TATACATACC | TTCTTCCAGG | GCATTAATGT | TCAAATGAAT | CATTTCCGTC | 7200 |
| | GCACGCATAT | TCATTTGATT | GTCCAACGCT | AAGTACGGAT | TAAAATGCTT | TGCATCATAT | 7260 |
| | AACAATATTT | GAAAATTTGA | ATCAGTCCCC | GTGACAATAC | ATGTATCATC | AGAATACAAA | 7320 |
| 20 | ATATTGCTTG | TTAATTTATT | AAATAGCAAT | GCCGTGAAAT | AGACCGGACG | TTTTCCATTA | 7380 |
| | TATTGATGAA | ATAGTTCAAT | AGAATTCATA | TAATCCCGTT | CATTTTTACA | ATGACTGACG | 7440 |
| | TGCAAATCAT | AATTCAACCA | ATACCCGATA | CCCTCTACTT | TAGAACTTAA | TTTTAATAAT | 7500 |
| 25 | TGCTCAATGA | TGATACCACC | TCTAAAATAT | TCGCCGTTTG | TAATAAATGT | ATCACCCGTC | 7560 |
| | AATGTATTCC | AATTGAGTAA | AATGAGTGGA | CGCTTTAGGC | GATGACGATG | CATTAAGTCG | 7620 |
| 30 | ATAAGGTAAT | TCGTTTTATT | AATAATCATT | TGACTCGCGG | TTTTAAATTC | ATCATCATTC | 7680 |
| | ATTTTATTAA | AATCAACAGC | GTCATTTGAA | TTGGCATÇAA | ATACAAAATG | GTCGATGTGT | 7740 |
| | GGCTCAAGTC | GTTTCAATAA | TGGTAGATGT | CTTTCCGTAG | CTTGATCTAA | GTGAATGTAC | 7800 |
| 35 | AAGCCACCAT | TAGGGAATAA | TGCTTTAAAA | TAATCAATCA | TTTCAATCAA | AGACGTGTGC | 7860 |
| | AATGTCGTCA | CATACAAGTT | GAACTTCAAA | TCTTTTCTAT | GACTGACATG | CAGGGCAACG | 7920 |
| | TGATGGATAA | AAATTTTAAA | TGCATCGATA | TAATCACGTG | AGTCATACTG | ATCCAAATGC | 7980 |
| 40 | ATGGTCAAAC | TAAAGTTATG | ATCTAATAAA | AAGTCTAAAC | ACAAATCAAT | ATCATAAAAT | 8040 |
| | ATATTCGAAA | TTTCTGCATC | ATACGTGAAT | GGCGCATTGA | GCTTTTTCAT | GATATATGGA | 8100 |
| 45 | ATCACATCAT | ATGCTAATAC | TTCATTGACT | TGAAAATCAT | GATGACATGT | AAGCAACTGT | 8160 |
| 45 | GATTGATACT | GTGTATTGAG | CAAATTCCTC | AAATAGCCCA | CTTGAATAAT | ATGATTAAAT | 8220 |
| | TGATTTAGTT | GGTGATTGGT | TGGTTGAAAG | GCAATCTCTT | TATAGTTCAT | CTTTTCAATA | 8280 |
| 50 | TCTTCAATAA | AATGATTCAT | TTCTTCAATG | TAGTCATTTA | AAAGTAATAT | CAATTCACGG | 8340 |
| | TCGTGATAAT | CATGTTGTGC | CGATTGCTGG | TTTTCAGTGA | TTGCTGGACG | ATCACCTCGA | 8400 |
| | TATTGTTTAG | GTGTTTGATG | CGTAAAGTGT | TTAAATGTTC | TCGCAAAGCT | CGCTGCACTT | 8460 |

| | TTCGCATGCT | CAATTCGCGT | CGTATTTAAG | AAATGATGGA | ATCCTACACC | TAGCGATTCT | 8580 |
|----|------------|------------|------------|------------|------------|------------|-------|
| | GTAAACTTTT | TAGACAGATG | GCTCTCTGAC | CACCCAACGT | ATTCGCTTAA | TTCTGAAAGG | 8640 |
| 5 | CTTAAATCTT | CATGAAAATG | TAACTCGATA | TAGTCGCATA | CTTGATTCAC | TTTATCATCA | 8700 |
| | TTTAAGATAC | TTTGGTTCGA | ATGATATGTA | CGCGGGACAT | AATGAATCAT | ATGCATAAGC | 8760 |
| 0 | AACTGAATCA | CAAGTTGTTG | CTCAGTCAAT | TTAGACAACT | CATTATGTCG | GATATGTGTT | 8820 |
| • | GAAACCAGTC | TTGCCATTAT | ATTTCTCAGT | TGATGTATAT | TCTTTGTTGT | GGTCGCATCT | 8880 |
| | GTTAAGTGAA | AATATAGACA | ATGCACATCA | TCAAACTTGT | CTGCTAAATA | TTTCATTTGG | 8940 |
| 5 | AATTGGATAT | AACATATGAT | GCCATCTTGT | TGAAGTTGAA | ATCGATACAA | GTCGCGGTGG | 9000 |
| | TTAATGATGA | AAATGTCGCC | ACTGTTGCAT | TGCGTCATAT | TATTTTCATC | ATAAATGTGT | 9060 |
| | GCCTCnCCTT | TAATAACAAA | ACCAATCATT | AAACTATTGA | GCCTTTTGAA | ATCTGACATA | 9120 |
| 20 | CTCTCAGTTT | CTACTCGAAT | TAAATAATCA | CGTTGCATAC | TATCCCTCAA | TTCAGTAATA | 9180 |
| | TGAATACGTT | TATTTTACAT | TATTTTACAG | CAACATATTT | GAATTTCATA | TTGAATCGTG | 9240 |
| | TGTGTGGATG | ATTATTTATC | CTCACTCGGT | TCAAGATGTA | GACTATCAGT | AAAAAAGTA | 9300 |
| ?5 | TTTTCACCTT | TTTTCTCCAC | AAAAGTAAAT | TCAATGTCTT | TATATCCAAC | TGrTGaACCT | 9360 |
| | TTTAAGTCTC | CCgAACCTTT | Caacaataac | TTTGGTGCTT | TATTCGTTGG | TATTTTATAT | 9420 |
| 30 | CTTTTTCGTA | ATTGTTTTAC | ATTATAGTCA | TCATTAGTTA | ATTGATATTT | TGCTGAATAA | 9480 |
| | CTCGGTACCT | CTGGATTATA | TGATATATCG | CCGTCTTTGT | ACTTCGACAA | ATCTTTAAAG | 9540 |
| | CTGCCATATT | GCGCGAAGAA | CTTAAAATTC | TCGATTTCTT | TTTTTATATT | TTCGTCTTTG | 9600 |
| 35 | ATACCTTTAG | TTGGAATGAT | TTTATTGTCT | ACCATTTTAA | CGGGATATTC | TTTATCTTTA | 9660 |
| | CTCTTAGGTC | TACCATCTTC | ATCATGAAGT | GTTTCACTCA | CTATATACTT | CCCGGTTGTA | 9720 |
| | GTCTŢĀGTGT | TTCTATTCAT | ATATAGAACC | ATACCTTTTG | ATTTCATACG | TTCCCCTTTA | 9780 |
| 10 | GGTTGAACAA | CCATTTCAGA | ACCAATAATC | CATGTACCTT | TATCATTTTT | ATCAAATTCG | 9840 |
| | TCATCACGAT | AACCTTCTTT | ATCGTATAAA | TCCTCTAGAT | TTTTAATCGG | ATACATACTC | 9900 |
| | AATGTTTTTT | CAAAGCTTTT | CTTAACTTCC | GCTTCTTTAC | CTATGCCACA | ACCAGCAGTG | 9960 |
| 15 | AAACTAATGA | CTAATATCAA | AAAACTAATA | TACAATACCA | ATTTGTTTAA | TCGTTTCATA | 10020 |
| | ATTTCACAAT | CCTATTCTTC | TTATTATCTT | TCCTGGATTG | ATTTCATATT | TTGATCGAGT | 10080 |
| 50 | CATGATTATT | TATCCTCACT | TGGTTTAAAA | ATTAACCCAT | CACTAAAGTA | AATGTTCTCT | 10140 |
| | TCTTTTTTCT | CTACAAACGT | AAATTCAATG | TCTTTATATC | CAACTGATGA | ACCTTTTAAA | 10200 |
| | TTCCCTGTAC | CVTTCAACAA | CArCTTCGGV | GCTTTATTTG | TTGGTATGTC | ATATCTTTTA | 10260 |

| ACCTCTGGAT | TATATGATAT | ATCTCCATCT | TTATAATTCA | TTAAATCTTT | AAAATTGCTA | 10380 |
|------------|------------|------------|------------|------------|------------|-------|
| TATTGCGCAA | AAAACTTAAA | GTTTTCGATT | TCTTTTTTTA | TGTLTTCTTC | TTTAACTTCC | 10440 |
| TCAGTAGAAA | TGAATTTATT | ATTAATCATT | TTAACTGGAT | ATTTTTTTTG | ATTATCCTGA | 10500 |
| GCTACTTCGT | ATTTCTCCGT | CTTTALTTCA | TTAGTATAGT | AAAAtCCTTT | TGCACTTCTT | 10560 |
| GTATTTCTAT | CTATCTTCAA | AAGCATGCCT | TTTATTTTTA | GAGCTTCTCC | TTTATTTTGA | 10620 |
| ATTGCCATTT | GAGAATTTAC | AATCCATGTT | CCCTTATCAT | TTTTATCAAA | TTGATCATCA | 10680 |
| CGATATCCTT | CTTTATCGTA | TAAATCCTCT | AGATTTTTAA | TCGGATACAT | ACTCAATGTT | 10740 |
| TTTTCAAAAC | TTTTCTTT | | | | | 10758 |

(2) INFORMATION FOR SEQ ID NO: 222:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1109 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

25

30

35 ·

40

45

50

10

15

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 222:

| 60 | TTCACAACCC | ACTITITGIG | TTTTTTAAAA | AgTCAATAAC | TGGaATTTAA | nTaTCaACTT |
|-------|------------|------------|------------|------------|------------|------------|
| 120 | CATTCCAAAA | ATTTTACCAG | ACAAGAACTT | ATTGCTTAAC | CAACGCGTTT | GCTTCTTTTT |
| 180 | TATATTATAC | ACGTTTAGCA | GTAATTTTGT | CAAAATAAAA | TAAAAACGTA | CAAATCAACA |
| 240 | TACCTAGACG | CTATAAATTT | TGCAATCGAG | TAACTTTTTG | GTAGCAGCTA | CTATTTATTT |
| 300 | GTTGCTCTAA | TTCCAAGATG | ATCTTTTGGA | ACGGTGCAAA | TGATATATTG | ATCATCTGAT |
| 360 | CGCCACCTTT | AACTTAGTAC | TTCATCAGCT | GAGTATTAAG | CCAAGTAATT | AGGTAATTCC |
| 420 | TGTTTTCAAT | AAATAACTCA | TTTACTTTCA | TACCCGTCTC | TATTCTTTAT | GCCAAAGACA |
| - 480 | CTGCAACAAA | CCTGCGCGAg | TGCCATCGCA | CCGTATGTTT | ATAGAATGAT | TACGCCAAGA |
| 540 | ACATCGTATG | CTTGAAGGTA | AATTTCCTTA | TCGTTACAAT | GGATGAGGTG | TGCTGCTGTA |
| 600 | ATTCAATGTC | AGTATTAAAT | TGGAAGATCG | CTGTTCCAGG | GCTACATCTC | AACATCTAAA |
| 660 | GCCCTCTCCA | CCTAACATTG | CAACATTTTA | AGAAATTCGT | ACTTCTGTAA | TCCCCATTTA |
| 720 | CGCCATGACG | ATAACTTTAA | GGCCATTGAT | CCACAAAAAA | GCATTTTCTT | TATAACTGGC |
| 780 | CCATCATATC | TCATCAATAC | TCCAGGCTTT | TCCCTTTAAT | ATTACTTCCT | TTCAACTGGA |
| 840 | CTTCACGAGC | ACTITITICC | TACTAATCCG | TATCGGCATC | AATCCATATA | TGGTACACTA |
| 900 | TACCGGAGGC | ACACCACCTT | AGATTTACCG | CTGCAACAGT | GCTAAATTTA | TAAGGCAACG |

| ATTITCTTCT | TTTGGTTTAA | ATTGATTTAC | TTTTCTTCC | GGCAATGTTT | CAAATCGTAT | 1020 |
|------------|------------|------------|------------|------------|------------|------|
| ACCGACCGTT | TTCGCACCGT | TTTCTTTTAA | TGCATTAACA | ACAGCCATCT | GTAAATCTAA | 1080 |
| aTTGCGtGCA | CCACCTAATT | GTGCCATTG | | | | 1109 |

(2) INFORMATION FOR SEQ ID NO: 223:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3997 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

15

20

25

30

35

40

45

50

5

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 223:

| TCTTTATTTA | AAAAAATGAT | TGTCTAGTTT | GTATCTCTCT | GAAGATTTGG | CAATAAATAA | 60 |
|------------|------------|------------|------------|------------|------------|------|
| AAGCCGATAA | CCGTATAATG | ATTATCGACT | TAAAGTTTAT | GTGGCATTTT | TTACTTTTGT | 120 |
| AATTTCAGGT | GAGTTAGATG | ATTATTATCA | GATAGATTAT | TGCTTATAAT | CATATGATGT | 180 |
| TTGAATGATA | TCTTTGATTT | CACTGATTAG | TGCTTCTTTA | GGATTAGCAG | TTGTACATTG | 240 |
| ATCTTCAAAT | GCGAGCTCTG | CCATTCTATC | AATTGACTCA | TTTAATTCTT | CTTCAGACAC | 300 |
| ACCTTGTGAT | TTCAAATTCA | TTTCAATTCC | GACTGATTGA | CCTAATTCGT | AGACAGCTTT | 360 |
| AgCTAATGAT | TCTACGAGTG | CTTCTGTCGT | ATTACCTTTT | AATCCTAAGA | ATTTGGCAAT | 420 |
| ATCTGCATAA | TCTGTATCTG | CTCTGAAGAA | CTCATATTTA | GGGAATAATG | CATGTTTTTG | 480 |
| CGGGTCTTTG | GCATTATAAC | GGATAATATG | CGGTAGTAAT | ATCGCATTCG | CTCTACCATG | 540 |
| CGGAATACCA | TATTCGCCAC | CAATTTTATG | CGCAATTGAG | TGTGCAATGC | CTAAGAATGC | 600 |
| ATTTGCAAAT | GCCATACCAG | CCAAAGTTGA | TGCGTTATGC | ATTTTCTCTC | TTGAAACTTT | 660 |
| ATCACCCTTT | TCAACAGATG | ATTTTAAATA | TTCGAACGTC | AATTTAATCG | CTTGTAGACT | 720 |
| CAAACCTCTT | GTGTAGTCTG | AAGCCATTAC | AGATACATAT | GATTCCATTG | CATGCGTTAG | 780 |
| TACATCCATT | CCTGTATCTG | CTGTAACGCT | TTTTGGCACA | CTCATCACAA | ATTGAGGGTC | 840 |
| AATAATTGCA | ACGTCAGGTG | TTAAAGCAAA | ATCAGCCAAC | GGATATTTTA | CATTTGTTTC | 900 |
| ACTATCTGTG | ATAACTGCAA | ATGGTGTTAC | TTCTGAACCT | GTACCTGATG | TCGTAGGGAT | 960 |
| ACAAATGAAC | GTCGCATTTT | CAGGCATGCC | TATTTTATAA | GTACGTTTAC | CGATGTCTAG | 1020 |
| GAACTTTTGT | TTAGCACCGA | AGAATGATGT | CTCAGGGTGT | TCAAAGAACA | TCCACATTGC | 1080 |
| TTTTGCAGCA | TCCATCGCTG | AACCACCACC | AAGTGCAATG | ATTGTATCCG | GTTGGAAATC | 1140 |
| AACCATCATT | TCCAGACCTT | TATATACTGT | ATTAGTTGAT | GGGTTCGGTT | CGACTTCGCT | 1200 |

| | ATAACCGAAT | TCTACCATAC | CAGGGTCACA | GACAATCATC | ACTITITCAA | TCTTGTCCAT | 1320 |
|----|------------|------------|------------|------------|------------|------------|------|
| _ | TGTTGTTAGA | CTCATGATTG | CATTTTCTTC | AAAATAAATT | TGAGCAGGCA | CCTTGAAAAT | 1380 |
| 5 | TTGAGTATTA | TTACGTCGTT | TAGCAATCGT | TTTAATGTTT | AATAAATCTG | TCGCACTAAC | 1440 |
| | ATTATGTGAA | ATTGAGTTTC | TACCGTAGaA | CCACAACCTA | ATGTTAAAGA | CGGAATCAAT | 1500 |
| 10 | TCGTTATACA | TATCACCAAT | ACCTCCAACC | GCTGATGGTG | TATTTACAAG | TACACGACAA | 1560 |
| | GCTTTCATTC | TTAGTCCAAA | ATCTTTTTGT | AATGTTTCAT | CTTCTGTATG | GATAACGGCT | 1620 |
| | GTGTGTCCTA | ATCCACCAAA | ATGTAGTGTG | TCTTCACAAA | TTTGAAATGC | TTGTTTTGTA | 1680 |
| 15 | GATTGGGCTT | TTACTAAGGC | TAATACTGGA | GATAATTTTT | CACGAGATAA | CGGATAGTCT | 1740 |
| | GAACCTACAC | CGCTAATTTC | GGCTATGATA | AGTTTTGTAT | TTTCGGGGAC | AGGTATACCT | 1800 |
| | GCTAATTCAG | CTATTTCAAC | TGCAGATTTA | CCGACAATAT | CAGGCTTAAT | ACCTGTTTTT | 1860 |
| 20 | TGTTCATTCA | TAATTGCATT | TTCTAAGCGT | TGTAATTCAT | CTTTTTTAAC | AAAGTATGCT | 1920 |
| | TGATGTGCTT | TAAATTCATT | AGTAACATCT | TTATAAATTT | CTTTATCAAT | GACTACAACT | 1980 |
| 05 | TGTTCAGAAG | CACAAATCAT | ACCATTATCA | AATGTTTTTG | AACCAATGAT | ATCATTTACT | 2040 |
| 25 | GCACGTTTAA | TGTGTGCTGT | TTTTTCAATG | TAAGACGGCA | CGTTACCTGG | TCCCACACCT | 2100 |
| | AATGCCGGTT | TGCCAGTTGA | ATATGCAGAC | TTAACCATGC | CCGAACCACC | TGTTGCTAGA | 2160 |
| 30 | ACTAATGCAA | TACCTTTGTG | ATTCATTAAT | TGTTTTGTTG | CTTCGATAGA | AGGCACTTCA | 2220 |
| | ATCCACTGAA | TAATATCTTT | AGGTGCACCT | GCCTTCATTG | CCGCTTCTAA | TACAACTTCT | 2280 |
| | GCTGCACGCT | TCGACGATTC | TTGTGCACTT | GGATGGAATG | CAAAAATGAT | TGGATTTCCT | 2340 |
| 35 | GTCTTAATTG | CAATCATCGC | TTTAAAAATA | GTTGTCGACG | TAGGATTTGT | TGTTGGCGTA | 2400 |
| | ACACCACAAA | TAACACCAAT | TGGTTCCGCT | ACATACGTTA | ATCCTTTTTC | TTTATCTTCA | 2460 |
| | CCAATAATCC | CTACTGTCTT | ATTGTCTTTT | ATTGAATTCC | ATATATATTC | AGAAGCGTAT | 2520 |
| 40 | AAATTTTTAA | TCGCTTTATC | TTCGTATATA | CCTCTTCCAG | TTTCTTCATG | TGCTAATTTT | 2580 |
| | GCTAGCACCA | TATGTTGATC | AACAGCTGCT | AAGCTCaTTT | GATGAACAAT | ATGATCAATT | 2640 |
| 45 | TCTTCTTGTG | ACTTTTTAGA | TAATGCTTCT | AATGCTTTTT | TCCCTTTGTC | AGCTAGAGCA | 2700 |
| 45 | TCAATCATAA | TTGCCACTTC | TTGTTCTTTC | GATCCACGAT | TTTCTTTTTC | AGGTATAGTT | 2760 |
| | AACATATACA | ACCACTCCTT | TATACTTTGT | GAATTATTTC | ACAAACATTA | TAGTACATGT | 2820 |
| 50 | CTCTCAGGAT | ATAAAGAAAA | TTCTATACAA | AAAAGTTTAA | TTTCGAATAT | TATTTGAACA | 2880 |
| | AATATCAAAT | TTTAAAATAA | ATGTTTTCAT | GAAATCATTG | TTATTTCGGT | GTTTTTAGAA | 2940 |
| | TGATTTTATA | ATCATAATTT | TTTCAATGAC | TTATTTAATA | CATAATTATA | TATTTAATTC | 3000 |

| | TCCTTGTCGA TACCTATCAA CAGATGTTAC AAATAAAAAC CaCCCGTGTG AACGGGTGGT | 3120 |
|----|--|------|
| | TTGTTCTGCG GCTATAAGCC TTCCTTACTG GCCAGCCCTA AAAGGGCACT GACAAGTCAG | 3180 |
| 5 | CCAACTGCAC TACTATTCCA GCAATCCTAA AGGTTTACTC TTTTTTCTTT CTTTTTTTAT | 3240 |
| | TTTTCTCTCC AGTGAAAGGA TCTAAATATT CTTCCATTGA AATTTCATCT GCAACGATAT | 3300 |
| 10 | CCTCTTGTAA TTGATTACGA ATATAATTTT CAATCACTTT TTTATTTCTA CCTACTGTAT | 3360 |
| | CCACATAAAA TCCTTTACAC CAAAACTTTC TATTTCCATA TCTATACTTT AAGTTAGCAT | 3420 |
| | GTCTATCAAA TATCATTAAA TTACTTTTCC TTTTAAATAG CCAACAAATG ATGATACCCC | 3480 |
| 15 | AAGTTTGGGT GGTATACTTA CTAACATATG GATATGATCT TTACATACCT CTGCTTCAAT | 3540 |
| | TATCTCTACA CCTTTTCTTT CATATAATTG ACGTAATATA ATCCCTATAT CTTTTTTTAT | 3600 |
| | TTTTCCATAT ATCGCTTGTC TTCTGTATTT AGGTGCAAAG ACAATATGTT ACTTACAATT | 3660 |
| 20 | CCATTTCGTA TGTGCTAAAC TGTTTGTGTC AGATGACATT AAATAGCATC TCCTCGTGTT | 3720 |
| | GATTATTTTG GTTGGCTGAC CAATATTTAC TCTAACATGT AGAGATGCAT TTTTTTGACA | 3780 |
| | ATGGTAGAAC CTTTTCTGGG GAGTGGGACA GAAATGATAT TTTCGCAAAA TTTATTTCGC | 3840 |
| 25 | CGTCCCACCC CAACTTGCAT TGTCTGTAGA AATTGGGAAT CCAATTTCTC TTTGTTGGGG | 3900 |
| | CCCCGCCCCA ACTCGCATTG CCTGTAGAAT TTCTTTTCGA AATTCTCTGT GTTGGGGCCC | 3960 |
| 30 | CTGACTAGAA TTGAAAAAAG CTTATTACAA GCGCATT | 3997 |
| 30 | (2) INFORMATION FOR SEQ ID NO: 224: | |
| 35 | (i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 1391 base pairs(B) TYPE: nucleic acid | |
| | $\bullet = \cdot$ | |

(C) STRANDEDNESS: double (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 224:

| GIIGCGNGNCN | MCACACLIA | IIGGIGCCAI | TATMCCTAGA | ATGAATTCAT | ATGCAGTAGA | 60 |
|-------------|------------|------------|------------|------------|--------------------|-----|
| TGABACAATC | AAAGGATTGG | CAAAACAATG | CCAAAAATAT | GAATCaCAAT | TAATTTTAAA | 120 |
| TTACACAGGT | TTAAATATCG | AAGCAGAAAT | ACAAGCGCTT | GAAACATTAG | CACGCAGTAA | 180 |
| AGTAGATGGT | ATTGTTTTAA | TGGCTACAGA | CATAACAGAG | AGACATATTG | AAGTCATTAA | 240 |
| TAAAATGAAT | GTACCAATCG | TTATTGTTGG | TCAACAACAT | GAACAACTTC | ATAGTATTGT | 300 |
| GCATGATGAT | TATAAAGCAG | GTCAAATTAT | AGGCGAATGG | ATTGGTCAAC | AGGGATATCA | 360 |
| ACAAGTTGAA | GTGTTTAGTG | ТААСТСАААА | AGATATTCCA | СТТССТАТАС | 5 T3C3333CC | 420 |

55

50

40

| | TACTTATGTG | GAAGCACAAA | AAGATGTTGC | AAATGTTTTG | GAAAATGTGG | AGCAAGTAGA | 540 |
|-----------|------------|------------|------------|------------|------------|------------|------|
| _ | TGCGGTTGTT | GGAGCAACTG | ATACGATTGC | ATTAGCTGCC | TATAAATATT | ATTCTGATAA | 600 |
| 5 | AAAAGATGTT | ATGAAACCAC | ATCAAATATA | TGGTTTTGGT | GGTGACCCAA | TGACACAATT | 660 |
| | AGTGTCTCCA | TCGATAAAAA | CAATTCATTA | TAATTATTTT | GAAGCTGGCC | AATGCGCGAT | 720 |
| 10 | GGaAGAGATA | CAACAGATGC | TTAAAAAGCA | AGATATGCCA | TATAGCGTCA | CAGTAGATGT | 780 |
| | TAATATTTAG | ACGCTGTATT | TTTTAAAATA | AATGTGGAAC | CGATACCATA | ТААСТАТААА | 840 |
| | TGGATAGGTT | AAAAGTTAAA | GAACGTAGGT | AAAATTTGCT | ATAATAGAAT | ATAAATTGTT | 900 |
| 15 | AACAGCATAA | ATTATAAAAG | GAGGACTGGG | TAAATATTAT | GACCGAATGG | ACTAGAGAAG | 960 |
| | AACGTTATCA | ACGAATCGAG | GACGTTGATA | CTGAGTATTT | TAAAACATTA | AAACAACAAG | 1020 |
| | TTGATCAATC | AAAATTTCGT | CAACAATTTC | ATATACAACC | AGAAACAGGC | TTATTAAATG | 1080 |
| 20 | ACCCCAACGG | ACTTATTTTT | TATAAAGGGA | AGTATTATGT | TTCACATCAA | TGGTTCCCAT | 1140 |
| | TAGGCGCAGT | ACATGGCTTA | AAGTATTGGT | ATAACTACAC | GAGTGATGAC | TTAATAAACT | 1200 |
| | TTAAAGCTGA | AGGGCCAATT | TTAAATCCAG | АТАСТАААТА | TGACAGCCAT | GGTGTATATA | 1260 |
| 25 | GCGGTAGCGC | TTTTGAATAT | AACGGGCATT | TATATTATAT | GTACACAĞGA | AATCATCGAG | 1320 |
| | ATAATCATTG | GCAACGACAT | GCGAGTACAG | ATGATCGCAC | GATTGAAAGA | AGACGGTTnC | 1380 |
| <i>30</i> | AGTTGGnAAA | A | | | | | 1391 |

(2) INFORMATION FOR SEQ ID NO: 225:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 930 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

40

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 225:

ATTTATTTA ATGTTATAT TTTCTAACAC TTTTTATGA TCATAGTAGT AATTGACATT 60
TTTCAATTCA AAGACTGGTG TCATCGTATC TCACCTCGCA TTCAACTATA CAACTCCTAG 120
TAACATATGT AAACAGTAAT GTTTACGACT CAAAATTAGA CAAAATAAAG AGATATGCCC 180
CCTTCAAGTT TTATTTATCG CATTTCTTGA AGAGAGCATT ATCATTTTAT TGTTGCATAA 240
CCTTATTTTT TAATTCTGGG TCAAATTGCT GTTGTTTAAA CATTTCAATT TCAAGTTTAT 300
ATGGCGGTTT TTTATTTTC TTATCTTCAC CAACATAAGG TGTTTCTAAG ATTTTCGGAA 360
TATCTTTAAA ACTATCATGA TGCACAATGT AATTTAATGC ATCAAAACCA ATGTAACCGA 420

55

45

| GAACAACTTT | GATTCTGTCG | ACTCCAATGA | TTTTATCAAA | TTCATTTAAT | ACGCCATCAA | 540 |
|------------|-------------|-------------|------------|------------|------------|-----|
| AGTCCTCTTT | AACATTATAT | CCAGCATCAT | GCGTATGACA | TGTATCAAAA | CATACTGATA | 600 |
| AACGTTCGTT | ATTATGAACT | CCATCAATAA | TACGTGCTAA | CTCTTCAAAT | GAGCGACCAA | 660 |
| TCTCTGTACC | TTTACCTGCC | ATCGTTTCAA | GCGCAATACG | TACATTATTG | TCATTCGTTA | 720 |
| AAACTTCATT | TAATCCTTCA | ATAATCTTAT | TAATTCCGGC | ATCAACACCA | GCTCCAACAT | 780 |
| GCGCACCTGG | ATGTAATACa | ATATCTTTAG | CCCCTATAGC | TTGCGTTCTk | TCaATTTCTT | 840 |
| GTTGCAAGAA | ATCTACACCA | AGATTAAACG | TTTCTGGTTT | GGTTGTALLG | CAATaTTaAT | 900 |
| GATGTATGGT | GGCATGAACA | ACAATATTAG | | | | 930 |
| (2) TYPODM | MTON BOD OF | 30 ID NO 30 | | | | |

(2) INFORMATION FOR SEQ ID NO: 226:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1984 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

25

30

35

40

5

10

15

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 226:

| TGACGCACCA | ATTTATAACG | CAATTGACAA | AACAATTAGA | TATACCTGTG | AAATTTGTAC | 60 |
|---------------------|------------|------------|------------|------------|------------|-----|
| CTGGAAACCA | TGATTTATGG | GAAGTTGAAA | GTATGACTAC | GCAAGACATT | TGGAATAATT | 120 |
| ATAAGAGTAT | GTCACAGTGC | TTGGTAGGAA | AACCATTTAT | agtaaatgaa | GAATGGGCAA | 180 |
| TCATAGGACA | TACTGGCTGG | TATGATTATA | GCTTTGCAGC | ACAACGATTT | TCATTAGATG | 240 |
| AGTTACAAAA | AGGAAAACAT | TATGGTGCGA | CTTGGCAAGA | TAAAGAACGA | ATATCTTGGG | 300 |
| GCATATCAGA | TCAAAATTTA | TCTAAAATAG | CGGCTGAACA | AGTGAAGAAA | GATATATTAG | 360 |
| aagt <u>ä</u> ggaaa | TAGACGAGTG | ATTTTAGTCA | CACATGTTGT | GACGCACCCT | GATTTCATTG | 420 |
| TTCCTATGCC | GCATCGTATA | TTCGATTTTT | ATAATGCATT | TATTGGGACA | AGTGATTTCA | 480 |
| ATCCTTTGTA | TGCGATGTTC | GATATACCAT | ATAGTATTAT | GGGCCATGTT | CATTTTCGTA | 540 |
| AAAGTGTGAT | AGATGATGGC | AGATGTTATC | TCTGTCCGTG | TCTAGGCTAT | CCAAGACAAT | 600 |
| GGCGTTCAGA | AGATATTTAC | CAGGAAATAA | ATGAGACGAT | ACAAATAATA | GAAATTTAAA | 660 |
| ATGCGCAAAC | CTGACCCAGT | TTGCGCATTT | TATGTTTTAC | ACACGCGAGT | AATGTGTTTA | 720 |
| CTTACGTGTG | TTTATTTTGT | TGCTGATTTT | CAATTGTATA | TGAATGTGGT | TGCACATAAA | 780 |
| TGCACTTTCT | TCCTGGTGAA | TTAAAGCTGT | ATTCCATTTT | CTCTTTACGA | ATTTTAATAA | 840 |
| TTTGTTTGCG | ATTTGGAATG | ATGGCAGGTA | AAACTAGGCC | ACGACGAATA | TGACTCCAAA | 900 |

55

| | TIGAAACTIG | TTTCGCTGGC | TTGTTATCAA | AGCGGAAAAC | ACGTAGTAAT | GGTTTAGAAC | 1020 |
|-------------|------------|------------|------------|------------|------------|------------|------|
| | CAAGATTAGT | ATGGTATATT | AACACAGGTT | GACCTTGATC | GATAATACCT | TTAAGATCTT | 1080 |
| 5 | CTAACGATTT | ACCAGTGCCG | TCTACGATAT | TAGGATTGTA | TTTTTGTAAA | AATGGTACAT | 1140 |
| | ATGCTTCTGG | AAATATCGTT | TGATGATAAT | TGCCAAGCTT | AATGAATAAG | TGATGTCCAA | 1200 |
| 10 | CATAACCTTT | ATGTGGATTG | TTCGGATGTG | TCGGCCAATG | TCTCATAATT | TCTGTAGCAG | 1260 |
| ,,, | GGATATGTTG | GTTGTTGTAT | TGCAACATCA | TGGCTGCGGA | AACACCTTCA | CACCCCATGA | 1320 |
| | CCATAGGGAT | AGGAAATAGC | TGACTGATAG | GTTTAACTGG | TAATATTTT | CGGTTCATAA | 1380 |
| 15 | TATAGTCCTC | GCATTGATTC | AATAAATATT | TAATATAATT | ATATAGCGTC | AATGCAAAAT | 1440 |
| | GTCCTAAACA | TATGTTTTAC | ATGAGTGAAT | AAAATTAATG | GAGTGATAAA | ATGGAATATC | 1500 |
| | AATTACAACA | ACTTGCGTCG | TTAACGTTAG | TAGGTATTAA | AGAAACGTAT | GAAAATGGAC | 1560 |
| 20 | GACAGGCTCA | GCAACATATA | GCAGGGTTTT | GGCAAAGATG | TTATCAAGAG | GGAGTAATTG | 1620 |
| | CGGATTTACA | GTTAAAAAAT | AATGGTGATT | TAGCCGGGAT | ACTTGGCTTA | TGTATACCTG | 1680 |
| | AATTAGACGG | TAAGATGTCA | TATATGATTG | CAGTTACCGG | AGATAATAGT | GCTGATATTG | 1740 |
| 25 | AAAAATATGA | TGTCATAACA | TTAGCAAGTT | CAAAGTATAT | GGTATTTGAA | GCACAGGGCG | 1800 |
| | CAGTACCTAA | AGCAGTTCAA | CAAAAAATGG | AAGAGGTTCA | TCACTACATA | CATCAATATC | 1860 |
| <i>30</i> . | AAGCAGATAC | GGTAAAATCA | GCACCATTTT | TTGAGTTGTA | TCAGGATGGT | GATACTACAA | 1920 |
| | GTGGAAAATT | AATATTACCA | GAAATTTGGG | AThCCTGTTA | AAGGGGTGAT | TGAAATAnGA | 1980 |
| | AnTG | | | • | | | 1984 |

(2) INFORMATION FOR SEQ ID NO: 227:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6373 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 227:

GATTCCACGT GTGTTAAAAG AAGTTACACC TTCAATGATG GTATTTACTA ATTTCTTTAG 60

AGATCAAATG GATCGCTTCG GTGAAATTGA TATTATGGTT AATAACATTG CAGAGACAAT 120

TAGTAATAAA GGCATCAAAT TATTGCTAAA TGCTGATGAT CCATTTGTGA GTCGTTTGAA 180

AATCGCAAGT GATACGATTG TGTACTATGG TATGAAAGCA CATGCCCATG AATTTGAACA 240

AAGTACGATG AATGAAAGTA GATATTGTCC AAACTGTGGT CGCTTATTGC AATACGATTA 300

55

35

40

45

| | AAAATATGAA | ATATCAAGTT | TTGATGTGGC | ACCGTTTTTA | TATTTAAATA | TCAATGATGA | 420 |
|----|------------|------------|------------|------------|------------------|------------|------|
| | AAAATATGAT | ATGAAAATTG | CAGGTGACTT | TAACGCTTAT | AACGCGTTAC | AGCATATACT | 480 |
| 5 | GTTTTAAGAG | AGCTAGGGTT | AAATGAACAA | ACAATTAAAA | ATGGCTTTGA | AACGTATACA | 540 |
| | TCAGACAATG | GTCGTATGCA | GTACTTTAAA | AAAGAACGAA | AAGAAGCGAT | GATCAATTTA | 600 |
| 10 | GCTAAAAATC | CTGCAGGAAT | GAATGCAAGT | TTATCAGTTG | GTGAACAATT | AGAAGGCGAA | 660 |
| | AAAGTGTATG | TTATTTCGCT | AAATGATAAC | GCTGCAGATG | GTCGAGATAC | TTCATGGATT | 720 |
| | TATGATGCAG | ATTTTGAAAA | ATTATCTAAG | CAACAAATTG | AAGCTATCAT | CGTGACAGGT | 780 |
| 15 | ACACGAGCAG | AAGAACTTCA | ATTGCGATTG | AAGTTAGCAG | AGGTTGAAGT | ACCAATTATA | 840 |
| | GTTGAGCGTG | ATATTTATAA | AGCAACGGCA | AAGACTATGG | ATTATAAAGG | TTTCACAGTT | 900 |
| | GCAATACCAA | ACTATACATC | ATTAGCGCCT | ATGCTTGAAC | AATTAAACCG | TTCGTTTGAA | 960 |
| 20 | GGAGGTCAAT | CATAATATGC | ATGAATTGAC | TATTTATCAT | TTTATGTCAG | ATAAATTGAA | 1020 |
| | TTTATACAGT | GATATAGGAA | ATATTATTGC | TTTAAGACAA | CGTGCTAAAA | AACGAAATAT | 1080 |
| | TAAAGTTAAT | GTCGTAGAAA | TCAATGAAAC | AGAAGGTATT | ACCTTTGATG | AATGTGATAT | 1140 |
| 25 | TTTCTTTATC | GGTGGTGGAA | GTGATAGAGA | ACAAGCATTA | GCAACAAAAG | AATTAAGTAA | 1200 |
| | AATTAAGACA | CCACTTAAAG | AAGCGATTGA | AGATGGTATG | CCGGGATTAA | CGATTTGTGG | 1260 |
| 30 | AGGCTATCAA | TTTTTAGGGA | TATATAAAAA | CACGCCTGAT | GGTACAGAAT | TAGAAGGGTT | 1320 |
| | AGGTATTTTA | GATTTTTATA | CTGAATCAAA | GACAAACCGA | TTAACAGGAG | ATATTGTTAT | 1380 |
| | CGAAAGTGAT | ACTTTTGGAA | CTATTGTAGG | TTTTGAAAAT | CACGGTGGTA | GAACATATCA | 1440 |
| 35 | TGATTTCGGT | ACACTTGGTC | ATGTTACTTT | TGGTTATGGT | AATAATGATG | AAGATAAAAA | 1500 |
| | AGAAGGCATT | CATTATAAAA | ATTTATTAGG | TACTTATTTA | CATGGACCAA | TTTTACCTAA | 1560 |
| | AAATTACGAA | ATCACTGATT | ATCTGTTAGA | AAAAGCTTGT | GAACGTAAGG | GTATTCCGTT | 1620 |
| 40 | TGAGCCTAAA | GAAATAGATA | ATGAAGCGGA | AATACAAGCG | AAACAAGTAT | TAATAGACAG | 1680 |
| | AGCAAATAGA | CAGAAGAAAT | CTCGTTAACT | CTGAACATCG | CATCAATGGA | TTTAATATTG | 1740 |
| i. | ATAAACGATG | AAGTTTAGTA | ATTAATCATA | TATGTATAAA | CACACACATT | ATTTTGGATG | 1800 |
| 45 | GAAACAACCA | AATTGATGTG | TGTTTTETTG | TTCTAGTGAA | TAATTATTAT | ACAATGAGTA | 1860 |
| | TCTATCCTAG | AATTATCAAT | AGTAATGGTG | ATTATGCAAC | ATGAAAAAAT | GAATGATGAA | 1920 |
| 50 | AGGAATTTGA | CGATGAAGCC | TACTAAAGTG | ATATTAAAAG | ATGCATCTTA | TTTACATAGC | 1986 |
| | AAAACATCGA | TAACATTTAT | TTTAAAAGAT | GTAGTTATCG | AAGAAGATAA | TAAAATTTAT | 204 |
| | татттсаса | СТВСТССВСТ | TTCGAAGATC | AAGAAGTTAA | עישרע ע פארידע ע | CCACTCTTTC | 210 |

| | TTATAGAACC | TGATTTACAT | TTTACAATTA | TTGATTTTA | A TCAAGAACTO | CTTTGTATTT | 2220 |
|-----|--------------|------------|------------|------------|--------------------|------------|-------|
| | ATATTGATTT | TGATTCTGGT | TTAAGGCATI | CAAACATGGC | AACAGAATCT | GGTATTTCAT | 2280 |
| 5 | TAAGGATAAA | TGTTGCTAAA | TCAGATTTTA | CTAAATTTAT | TAATGAATTA | GCCTCTTTAC | 2340 |
| | ATTAATGATT | TAAATCTGAT | ATGTAATTAC | AATCAAAAA | GACAGCCACA | TCCCTCCGTA | 2400 |
| 10 | GTTTAGGCGT | GTGGCTATAT | TTGAGTCTGA | ATATTTATGO | TTGTAATTT | AAAAAGGGAC | 2460 |
| ,,, | ATGCTATATA | CGATAAAAAG | AGGCGGGGAC | ATAAATCAAT | GTTCTATGCT | CTACGAAGTT | 2520 |
| | ATATTGGCAG | TAGTTGACTG | AACGAAAATG | CGCTTGTAAC | AAGCTTTTTT | CAATTCTAGT | 2580 |
| 15 | CAGGGGCCCC | AACAAAGAGA | AATTGGATTC | CCAATTTCTA | CAGACAATGC | AAGTTGGGGT | 2640 |
| | GTGGGCCCCA | ACACAGAGAA | TTTCGAAAAG | AAATTCTACA | GGCAAAGCGA | GTTGGGGTGG | 2700 |
| | GACGACGAAA | TAAATTTTAT | GAAAATATCA | TTTCTGTCCC | ACTCCCATGG | TGCCAATTAG | 2760 |
| 20 | CATAAGGTAC | TTAAATTAAG | CATATCTGCT | GTCTAGCAGT | CGATAAATCA | TTAGAACTTC | 2820 |
| | GTATAGTATA | TGACTTTTAA | TTTGATTTTC | ACCACTAATT | TCAAGTGCTT | TTATAGTCGA | 2880 |
| | ACGTAAAGTT | TCTACAGAAT | CATCTTCTCT | CTTAAAAGAA | CCATCATAAA | ATATATCTTT | 2940 |
| 25 | GATGCTACTA | CTAATTTTTA | GCAATGCCAT | TTTTTCGTCA | CCTGAAAAGT | TAACACGAGT | 3000 |
| | ATTTTTAGGC | AAGTAAATGA | TATTTGATAA | ATGAGTGATA | AACAAACGAT | TCGTATATGC | 3060 |
| 30 | ACGTTTAGTT | AATTGATTGA | GTAATTTCCA | ATCACATTCT | TTTTTCTTAT | GATAGCTTAA | 3120 |
| | TTCATCACGT | TGATAACTTA | TTAACGTTTC | AACTTGATTA | TTTAAATTGA | AAATATTTTT | 3180 |
| | ATATGCTTTT | TCGCTTTTAT | CAGATTGCAG | TCTTGATAAG | ATAAGTTCTT | GGCAGCGATT | 3240 |
| 35 | GTAAAATAAT | TTATACATCA | AGGCATCTGT | CTTACTTAAT | TTTTCTTCGA | CCTGACCATA | 3300 |
| | ATACTTAGGT | GGAAACACCA | TGAAGTTAAT | TAAACCTGAT | GTCACGAGTC | CAATAATTGC | 3360 |
| | TGTCAATGTT | CGAGACAAAA | AGTTGAATAT | GTAGGCATCA | TGAATACCTG | GAATCATAGC | 3420 |
| 40 | TAATGATGTT | AGTACAGCGA | CATTCGTACC | AACTTGCAAT | TTGAGTTTTG | TACAGAATAA | 34.80 |
| | AATCGTGAAC (| GTTGCACTCA | ATGCATATGT | AAAAGGTGAT | TGATCGCCGA | ATAAATATGT | 3540 |
| 4.5 | AAATAATACT (| GCAAAGCCTG | CACCAATTAC | CGTAGCAGGT | AATCTACGAT | AACCTTTAAT | 3600 |
| 45 | AAGTGATGCC : | TTGGCAGTTG | GTTCAATTGT | GACTACAGCT | GTTAAAATGG | CATAGATGGG | 3660 |
| | TGTTAAATCT I | AGTGCCATAC | AAAAGACAGC | TGTTAAAAAA | ATGGCAATAC | CAGTTTTAAT | 3720 |
| 50 | TGTTCTGGCA (| CCAATTAAAT | GTTTATACCA | TTGATCGTTC | ATTTTTTAAC | CTCTAATCAT | 3780 |
| | CGTAAAATCT 1 | TAGCGAGCGC | TTTATAATAA | TAGTATCGTA | CATTGGAAAA | GTTCATGTAT | 3840 |
| | GTAAAATATT 1 | GAAATAATC | ATACATAAGC | ATTACTTTGA | ТТ ТТСАТАТА | САТТАВТСАВ | 2000 |

| | CAAGCATTTT | TCAATTATAG | TCCGGGGCCC | CAACATAGAG | AATTTCAAAA | AAGAAATTCT | 4020 |
|----|------------|------------|------------|------------|------------|------------|------|
| | ACAGGCAATG | CAGGTTGGCG | GGGCCCCAAC | ACAGAAGCTG | ACGAAAAGTC | AGCTTACGAT | 4080 |
| 5 | AATGTGCAGG | TTGGCGGGGC | CCCAACATAG | AGAAATTGGA | TCTACAATTT | CTACAGGCAA | 4140 |
| | TGCAAGTTGG | GGTACAACGA | TAAAGAAATA | TITTTTCTTT | ATCACACTAT | GTCTCACTCA | 4200 |
| 0 | CTTTCCAAAA | TACTAAAGTA | ACATCTTTAG | TATATCAAAG | AATTTTTGCT | ATAATAAGTT | 4260 |
| | ATAATTATAT | AAAAAAGGAA | CGGGATAAAA | TGATTGTAAA | AACAGAAGAA | GAATTACAAG | 4320 |
| | CGTTAAAAGA | AATTGGATAC | ATATGCGCTA | AAGTGCGCAA | TACAATGCAA | GCTGCAACCA | 4380 |
| 15 | AACCAGGTAT | CACTACGAAA | GAGCTTGATA | ATATTGCGAA | AGAGTTATTT | GAAGAATACG | 4440 |
| | GTGCTATTTC | TGCGCCAATT | CATGATGAAA | ATTTTCCTGG | TCAAACGTGT | ATTAGTGTCA | 4500 |
| | ATGAAGAGGT | GGCACATGGG | ATTCCAAGTA | AGCGTGTCAT | TCGTGAAGGA | GATTTAGTAA | 4560 |
| 20 | ATATTGATGT | ATCGGCTTTG | AAGAATGGCT | ATTATGCAGA | TACAGGCATT | TCATTTGTCG | 4620 |
| | TTGGAGAATC | AGATGATCCA | ATGAAACAAA | AAGTATGTGA | CGTAGCAACG | ATGGCATTTG | 4680 |
| | AGAATGCAAT | TGCAAAAGTA | AAACCGGGTA | CTAAGTTAAG | TAACATTGGT | AAAGCGGTGC | 4740 |
| ?5 | ATAATACAGC | TAGACAAAAT | GATTTGAAAG | TCATTAAAAA | CTTAACAGGT | CATGGTGTTG | 4800 |
| | GTTTATCATT | ACATGAAGCA | CCAGCACATG | TACTTAATTA | CTTTGATCCA | AAAGACAAAA | 4860 |
| 30 | CATTATTAAC | TGAAGGTATG | GTATTAGCTA | TTGAACCGTT | TATCTCATCA | AATGCATCAT | 4920 |
| | TTGTTACAGA | AGGTAAAAAT | GAATGGGCTT | TTGAAACGAĞ | CGATAAAAGT | TTTGTTGCTC | 4980 |
| | AAATTGAGCA | TACGGTTATC | GTGACTAAGG | ATGGTCCGAT | TTTAACGACA | AAGATTGAAG | 5040 |
| 35 | AAGAATAGTT | CAACATATAC | TAAGACTAAA | GTATGAACAT | CATTTAGTTC | CGGAGCCTAT | 5100 |
| | TCATATTGGT | TTCGGAACTG | TTTTATAATA | ATTAAGAACA | CAATCAATGC | GTCATTTCAA | 5160 |
| | AAATATGTTG | TAACAAAGTA | GTTTTTAAGC | AAACATATCA | TCGACATCAA | CGAAGATACA | 5220 |
| 40 | TAGCGCATTT | GGTATTTTAA | AACTTATTAT | AAAAGGTGAT | AGTTATGAAC | TATGTTGAAC | 5280 |
| | GTTATATTGA | ACAGTTTTTG | AGAGCAACAG | TAAGAAATAA | TATCAAGCAC | TACCTTTTAA | 5340 |
| | TGCTAGATGA | AAAAATGAAA | AATTTAGATG | ATTATATGCG | TTATTTAATT | ACTAAAAAAG | 5400 |
| 45 | AACAACTTAG | CAAGTTAATT | GACAGTCTAA | TGCTAACATT | AGAAAATAAA | TATATTGATA | 5460 |
| | TTGCTGAAGC | ATTTCAAATT | CAATGTGCAA | GAGAAATCAA | TAATCAAGAA | ATTGAAAATA | 5520 |
| 50 | TTAAATCAGA | GTTGAATAAA | GTTGAAGCAT | ATTATGCACA | AATTGAAACT | CAAATTCAAC | 558 |
| | AAACTTCAAC | TGAAAAAATA | GCAACAGAAA | AAACATCGTA | TCTAATAAAT | TATATGAACG | 564 |
| | CTGTGGCATA | GAAAGGCGGC | GAAACATGAC | ACACAAATAT | ATATCAACGC | AAATGTTGAT | 570 |

| CTTTTTACTC | GTTCTATTAT | TGGGATGTGT | ATTAGTTTAT | GTAGGATATC | TTTATTTTCA | 5820 |
|------------|------------|------------|------------|------------|------------|------|
| TAAAATACGT | GGCCTTTTGG | CGTTTTGGAT | AGGCGCGCTA | TTAATTGCAT | TCACATTATT | 5880 |
| GTCTAATAAG | TATACAATCA | TCATCTTGTT | CGTCTTTTTA | TTATTACTTA | TTGTGCGTTA | 5940 |
| TTTAATACAC | AAGTTTAAAC | CAAAAAAAGT | AGTTGCGACG | GATGAGGTTA | TGACTTCACC | 6000 |
| ATCTTTTATT | AAACAAAAGT | GGTTTGGTGA | GCAACGTACA | CCAGTTTATG | TATATAAGTG | 6060 |
| GGAAGATGTA | CAAATTCAAC | ATGGAATTGG | CGACCTACAT | ATTGACTTAA | CAAAAGCTGC | 6120 |
| AAATATTAAG | GAAAATAATA | CCATTGTTGT | TAGACACATT | TTAGGTAAAG | TGCAGGTTAT | 6180 |
| ATTGCCGGTT | AATTACAATA | TTAATTTACA | TGTAGCTGCT | TTTTATGGAA | GTACTTACGT | 6240 |
| GAATGÁAAAA | TCATATAAAG | TTGAAAATAA | CAATATTCAT | ATTGAAGAAA | TGATGAAACC | 6300 |
| GGATAACTAT | ACAGTTAATA | TCTACGTATC | AACGTTTATC | GGAGACGTAG | AGGTGATTYA | 6360 |
| TCGATGAAyC | ACT | | | | | 6373 |
| | | | | | | |

(2) INFORMATION FOR SEQ ID NO: 228:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4488 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

30

35

40

45

50

10

15

20

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 228:

| A | IAGIIGAAAG | CGITTIACAC | TTAATAACTC | CCTCTTAAAT | GCATCCAGGT | TCTATGTAGT | ` | 60 |
|----|------------|------------|------------|------------|------------|------------|---|-----|
| A | AATCATGAA | nATAACATAT | AAATnTAGAG | GAGATTTACC | TTTGAATACA | GAGAACAACA | | 120 |
| A | GAATCAAAA | CCAATCTGTT | AAAAATTCTG | AAAGaCGCGG | CATGTTAAAA | GGATGCGGCG | | 180 |
| G: | TTĢCCTTAT | TTCTTTTATT | TTATTAATAA | TCTTATTATC | AGCCTGTTCA | ATGATGTTTA | | 240 |
| G. | FAATAATGA | CAATTCCACT | AATAATCAAT | CATCAAAAAC | GCAATTAACT | CAAAAAGATG | | 300 |
| A | AAATAAAA | TGAAGATAAG | CCTGAGGAAA | AATCAGAAAC | AGCAACAGAT | GAGGATTTAC | | 360 |
| A | ATCAACCGA | AGAAGTACCT | GCAAATGAAA | ATACTGAAAA | TAATCAACAT | GAAATTGATG | | 420 |
| A | AATAACAAC | AAAAGATCAA | TCAGACGATG | ATATTAACAC | ACCAAACGTT | GCAGAAGATA | | 480 |
| AJ | ATCACAAGA | CGACTTGAAA | GATGATTTAA | AAGAAAAGCA | ACAATCAAGT | AACCATCATC | | 540 |
| A | ATCCACGCA | ACCTAAGACC | TCACCATCAA | CTGAAACAAA | CACGCAACAA | TCATTTGCTA | | 600 |
| A? | TTGTAAGCA | ACTTAGACAA | GTATATCCGA | ATGGTGTCAC | TGCCGATCAT | CCAGCATATC | | 660 |
| G/ | ACCACATTT | AGATAGAGAT | AAAGATAAAC | GTGCATGTGA | ACCTGATAAA | TATTABACAA | | 720 |

| | GGGAGATTTT | TTAGGCATGA | GCAATCAATT | CAAAAGCGAA | GAAGAGCGAA | GACAATGGGA | 840 |
|------------|-------------------|------------|------------|------------|------------|------------|------|
| | ACAATTCCAA | GCTTTCCAAA | ATCAACAAAA | CCAACAGAAC | CAGCAATACG | GACAAAAGAA | 900 |
| 5 | АТСТААААА | GGATGGTTCT | GGGGCTGTGG | TGGTTGTCTA | GTATTATTTA | TTTTAATTAT | 960 |
| | CATCGGTATT | TCAGCTTGTA | CAGCTGGTAT | TACAGGTAAC | CTTGGCGGAA | ATAGTTCTAA | 1020 |
| 10 | AGAAACGAAC | AAAACCCATA | AAATCGGTGa | AACTGTTAAA | AATGGCGACC | TTGAAGTCAC | 1080 |
| 10 | TGTAAATTCA | GTGGAAACTA | TGAAATCTGT | AGGACCATCT | CTTGCACCAA | CAAACGCTAA | 1140 |
| | AGGTATATTT | GTCGTTGCTG | ATGTGACGAT | TAAAAACAAA | GGTAAAGAAG | CGTTAACAAT | 1200 |
| 15 | TGATAGTTCA | ATGTTTAAGC | TAAAATCCGG | TGATAAAACA | TTTGAAGCAG | ATAATACAGG | 1260 |
| | TTCAATGTCT | GCTAATCAAA | GTGACAATGG | TAGTATAGAA | AATTCATTTT | TCTTACAGCG | 1320 |
| | TATAAATCCA | GATAGCACTG | CTCAAGGTAA | AATTGTTTcG | ATGTGTCAGA | AAACATAGCC | 1380 |
| 20 | AACGCAAAaG | ATAAAAAATT | AGAAGTTATT | TCTAGTTTAT | TTAGCGTCAA | GAAGATTACA | 1440 |
| | TTTGATTTAT | CCGATGCTAA | AAAAACATCA | AAAGCTAAAA | AAGACAAGCA | AGATACAGAA | 1500 |
| | GTAGCTGTTG | CGAGTTCAAA | TAGCGATAAT | GTAAGTTATG | AAGCTTCGGC | TACTACACCT | 1560 |
| 25 | GCTACAACTT | CTAGTGCGGA | TACTGATTCT | GAAGATAGCG | AAAAGTCTAG | TAAAGATGAG | 1620 |
| | GATAAGCAGA | ATGCGTCTAA | Aagtgataaa | TCTAGTGTAG | aaaaagtga | ATCTAATGAG | 1680 |
| 3 0 | GAAACTGCTC | CTGTAGAGCC | CATGCCCCAT | AGCAAACCTA | CCACTAGTGA | AGCACCACCT | 1740 |
| | AGCCAAAATA | TTCACAaTGa | AGATAGCmTG | TACGACGCTT | CAACAGAATA | AAATtnyCAG | 1800 |
| | LAGCTCGGCT | ACCCTTCTTT | TACGGAAAAA | TTAATTATAC | ATAATCaAAT | CaAGGAGATA | 1860 |
| 35 | AAAAAATGAA | ATTCAAAGCT | ATCGTTGCAA | TCACATTATC | aTTGTCACTA | TTAACTGCCT | 1920 |
| | GTGGTGCTAA | TCAACATAAA | GAAAATAGTA | GTAAATCAAA | TGACACTAAT | AAAAAGACGC | 1980 |
| | AACAÃACTGA | CAACACTACA | CAGTCAAATA | CAGAAAAGCA | AATGACACCA | CAAGAAGCCG | 2040 |
| 40 | AAGATATAGT | TCGAAACGAT | TACAAAGCAA | GAGGCGTTAA | TGAATATCAA | ACATTAAATT | 2100 |
| | ATAAAACAAA | TCTTGAACGA | AGCAATGAAC | ATGAATATTA | TGTTGAACAT | CTAGTCCGCG | 2160 |
| | ATGCAGTTGG | CACACCATTA | AAACGTTGTG | CTATTGTTAA | TCGACACAAT | GGCACAATTA | 2220 |
| 45 | TTAATATTTT | TGATGATATG | TCAGAAAAAG | ACAAAGAAGA | ATTTGAAGCA | TTTAAAAAGA | 2280 |
| | GAAGCCCTAA | ATACAATCCA | GGTATGAACA | ATCATGATGA | AACAGATGGT | GAGTCAGAAG | 2340 |
| 50 | ACATTCAACA | TCATGACATT | GATAATAACA | AAGCAATTCA | AAATGACATA | CCAGATCAAA | 2400 |
| | AAGTCGACGA | TAAAAATGAT | AAAAATGCTG | TTAATAAAGA | AGAAAAACAT | GATAATGGGG | 2460 |
| | C3.3.503.3000C | ma | | | | | |

| | ACTATGCATG | GTCTTTTTAA | TCAACTTAAA | CTCGGCATTA | TTTCAATCGA | AAACGCAGAG | 2640 |
|----|------------|------------|-------------|--------------|--------------|--------------|------|
| | CATACGCTTT | TTACACCTTA | TATGTTGGAA | ACGCTCTCTT | CCCTAGGCGT | GAAAGACAGC | 2700 |
| ; | ATTGTCGATT | TAATTCATAA | AGGGACTGAA | TTAGAAGACT | TTGCGGCATT | TAATTTATCA | 2760 |
| | ATTGAAGACA | CAGTTACAGT | CTGTTTACAA | AGAACTGAAG | AACTATTAAA | ACAATACAAA | 2820 |
| • | AATGTGGAAT | TCAATGACAA | AATATTAATC | AATTGGCGTA | TTATACAAGA | GAAATAGACA | 2880 |
| 0 | TATAAAAGTC | GAATGTAACt | ACGTGAGTAT | TGATTTTATT | CTTTGTAAAT | TACAAGCATT | 2940 |
| | TCATATTATA | AAGTTTGAAA | AGAGGTATAT | TGAAATGGAG | AAAAATGAAT | ATATAGCTAA | 3000 |
| 5 | ATATAATGAA | TATAGTCAAT | TATTAGACGC | TACATACTCG | CAAGCTGTAG | CATMCCTTTT | 3060 |
| | AAGtAAATaT | GGCGCTGTAA | CCGATGATTA | TTATAAAGaA | AAATCATACA | CGCGATTTTT | 3120 |
| | AAAtGGAGnA | ATCAAAAGTA | TTTCAAAAGG | AAAATACACT | AGAGCTAGCG | AaGGATTATA | 3180 |
| ro | TTGCCATCAT | ATAAGCGAGG | ACAAATTCCA | AAATCTATCT | GATCTAAGAT | TCATtTCCAA | 3240 |
| | ATTTAAGTAC | TCATACGACG | TTCAAAAGAA | AGAAAACTTA | GTGTACTGTG | ATCTAATCGA | 3300 |
| | GCATTTAATT | TTACATGCAA | TTATTACAAA | AGAATCCCAT | GGCCAATTTG | GTGTAGCTGG | 3360 |
| 25 | ATTATGTCAA | ATGATCAAAC | CAACAGTCAT | TGATTGGTAC | ATTGGCGAAT | ATAATCCAAA | 3420 |
| | ACCAGCATGG | ATGCAAGCCA | CCAAAGCACG | TGCCTATTTG | CCTGGAATAT | TAGTAGAGAA | 3480 |
| 30 | ATTACTCATT | AAAATTGACG | ATATGTTAAA | AGGAATAGAA | ATATAAGATT | TCCTTGAGTC | 3540 |
| ~ | TAGATAAATG | ATTAATGTAG | ATTTATTTTT | TGCTGTTGAG | ATTTTGTTAT | AGATGTTTAA | 3600 |
| | ACCTGTAATT | TTTATATAAA | DATAAAATAT | ACCACGCATA | CCTATCTATA | AACGGrCAAT | 3660 |
| 35 | GTTTATAAAT | GAGTTTGCAT | GGtCTTGAAT | TGTATTAAAT | TTCTTTTGGT | TTTAATAAAT | 3720 |
| | CGACTAGATT | TTCACAATAT | TTATCAAATA | TGTATTCCTA | AATTATACAG | CCTTAATCCA | 3780 |
| | GCAGCTACTT | TCGAAACTTC | CAACTTAGTT | GATATAAGGT | TCAATAGTTT | GTTTCGTTCT | 3840 |
| 40 | TTTTCAGATA | AACCAGAACT | TAAATTGATA | TTATTGACTI | CATAAAAATT | ATAGACTAAT | 3900 |
| | GCCTCTATTI | GCTTTTTAGG | CATAAGTAAG | TCGACTGAAA | ACTGATTTAC | GTCGCTTTCA | 3960 |
| | TAAATCATTT | CATGTAAATT | CTTTAGACTA | A TTATCGTTAC | TATCTCTCAT | TAAGTCTGTA | 4020 |
| 45 | TTTTTAAATA | AATAACGGCC | : CAATTCACG | a gctattgaaf | ATCTTGTATI | ATTAATCGAG | 4080 |
| | TGATTATTAT | TAATATAGAT | TGTTCTTCC | A CTTAAATAA | CCGAAGTAT | ACCCTCCATT | 414 |
| 50 | TTAATATAT | TAACATTTAA | ATTAAGTTG | A AATAATAGCT | TGTCTATGT | AATAGCAAAG | 420 |
| - | TGTTCAGAA | TAATAAAAA | TTGATCCAT | TTGTCCTTT | A TAAATGCCTC | G AAATAATCGA | 426 |
| | ACTATTTTT | ATTCTAAAAT | ATCTTCATA | A TGAACTTTC | CAATAACTT | CAATTGATTC | 432 |

| | TATATGTAAC GAAGGGACAT GATTTCAAAA | 4440 |
|------------|--|------|
| | TAAAATACCT TTTTTATAAA TNTATTATAA TATCCCCCAC TATACNAC | 4488 |
| 5 | (2) INFORMATION FOR SEQ ID NO: 229: | |
| 10 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 846 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| 15 | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 229: | |
| | TATGGCGCCA TATTAGTTGT AACTGGTTTA AGAGGTCCAA GAAATATCAA ATAAAGTTGT | 60 |
| | tCCTGGGCTT GGTACTGTTA TCTCAATATT GmwTGCaTTT GGTGGTCTAG CTTTTAATAT | 120 |
| 20 | TGGTAATATT GCTGGTGCCG GTTTAGGTTT AAATGCAATT TTTGGATTAG ATGTAAAATG | 180 |
| | GGGCGCAGCT ATTACTGCAA TCTTTGCAAT ATTAATCTTT GTAAGTAAAA GTGGCCAAAA | 240 |
| | AATTATGGAC GTTGTTTCAA TGATTCTTGG TATTGTGATG ATTTTAGTTG TGGCATATGT | 300 |
| 25 | GATGTTTGTT TCTAATCCAC CTTATGGTGA TGCTTTTGTG CATACATTTG CGCCAGAACA | 360 |
| | TCCAATGAAA TTAGTCTTGC CCATCATTAC GTTAGTTGGT GGAACTGTar GTGGTTATAT | 420 |
| 3 <i>0</i> | TACCTTTGCA GGTGCACATC GTATATTAGA CTCTGGCATT AAAGGTAAGC AATATTTACC | 480 |
| | ATTIGTAAAT CAATCAGCAA TIGCIGGTAT TITAACTACA GGTATTATGA GAACGTTACT | 540 |
| | ATTCCTAGCG GTATTAGGAG TTGTTGTAAC AGGTGTGACA CTAAGTTCTG AAAATCCACC | 600 |
| 35 | AGCGTCAGTT TTTGAACACG CAATTGGACC AATTGGAAAG AATATTTTTG GTATTGTGTT | 660 |
| | ATTTGCTGCA GCTATGTCAT CAGTAATTGG CTCAGCATAC ACAAGCGCAA CATTTTTAAA | 720 |
| | AACACTTCAT AAATCACTTA ACGAAAGAAG TAATTTAATT | 780 |
| 10 | TTCAACAATG ATTTTCTTAT TTATTGGAAA ACCAATCAGC CTTTTAATTA TAGCAGGCGC | 840 |
| | GATAAA | 846 |
| 4.5 | (2) INFORMATION FOR SEQ ID NO: 230: | |
| 15 50 | (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2072 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear | |
| | (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 230: | |

| | TCTTTTTAAA | AGGTACTAAT | ATTTCTTTAG | TGAAAATTGA | ATCACGGTCG | TTTATTGGTG | 120 |
|----|------------|-------------|------------|------------|------------|------------|-------|
| 5 | CCTTGAGTAT | ATTATTATAG | ACGGAATCTG | ATCTAATAAT | ATTGATTTTA | TACATGATAA | 180 |
| | ACCTCCTTAT | GTTGTCAGCA | TAAAGGATAA | CGTAACGTGA | TTTTCAAGCA | GTAATTGTAA | 240 |
| | CTAATTGAMA | AAAATTAAGA | AAAGTATGTG | AGTGTTCCTA | AwTAATATGa | TTAAAATGAT | 300 |
| o | GGCGAATAAG | TGTCTaAAAG | CATCTTAAAG | GGACATTGTA | TAGGGTAAAT | CACTTCATAA | 360 |
| | ATAAGGGAAA | ATCCTTATGT | TCACTTTTTC | ACAATCATnA | TAAAATATAT | ATGTAGTCAA | 420 |
| 5 | TACTTTGTCT | ATATTGAATG | TTTTCATATA | AATGAAAGCA | TTTTTAAATA | ACATTGACCT | 480 |
| | СТААТАТАТА | GGCAGAGTAT | TGATATCTAT | таааааатаа | ATGATTTTGA | TGAAGGTGAA | 540 |
| | ACGTATGTAC | AAAACAAAAG | GTGGCTTTCA | ACTTACATTA | CAAACATTAA | GTTTAGTGGT | 600 |
| | TGGGTTTATG | GCTTGGAGTA | TAATTGCGCC | ATTAATGCCC | TTTATTAAAC | AAGATGTCAA | 660 |
| 20 | TGTTACTGAA | GGTCAAATAT | CAATCATTTT | AGCGATACCA | GTTATTTTGG | GATCGGTGCT | 720 |
| | CCGTGTGCCA | TTTGGTTATT | TAACAAACAT | TGTTGGCGCT | AAATGGGTAT | TCTTTACTAG | 780 |
| ?5 | TTTTATCGTA | TTGTTATTCC | CGATATTTT | CTTAAGCCAA | GCACAAACAC | CGGGTATGTT | 840 |
| | AATGGCTTCA | GGATTTTTCC | TTGGTGTAGG | TGGTGCAATT | TTCTCAGTTG | GTGTTACATC | 900 |
| | AGTTCCTAAA | TATTTCCCTA | AAGAAAAAGT | AGGTCTAGCA | AATGGTATTT | ATGGTATGGG | 960 |
| 30 | AAATATCGGT | ACAGCAGTTT | CTTCATTTTT | AGCACCACCG | ATAGCGGGTA | TTATTGGTTG | 1020 |
| | GCAAACAACA | GTTAGAAGTT | ACTTAATTAT | TATCGCTTTA | TTTGCATTAA | TTATGTTCAT | 1080 |
| 35 | TTTTGGTGAC | ACACAAGAAC | GTAAAATTAA | AGTACCATTA | ATGGCtCAAA | TGAAAmCATT | 1140 |
| | ATCTAAAAAC | TACAAATTAT | ATTACTTAAG | TTATTGGTAT | TTTATTACTT | TTGGTGCTTT | 1200 |
| | TGTAGCATTT | GGTATTTTCT | TACCTAACTA | CTTAGTAAAT | CATTTTGGAA | TTGACAAAGT | 1260 |
| 40 | AGATÉCTGGT | ATTCGATCAG | GTGTATTCAT | TGCGCTGGCA | ACATTCTTAA | GACCAATAGG | 1320 |
| | TGGCATTTTA | GGTGATAAAT | TTAATGCAGT | TAAAGTATTG | ATGATTGATT | TTGTTGTTAT | 1380 |
| | GATTATCGGT | GCCATTATTT | TAGGTATTTC | AGACCATATC | GCATTATTCA | CTGTAGGCTG | 1440 |
| 45 | TTTAACAATA | AGTATTTGTG | CAGGTATTGG | TAACGGCTTA | ATCTTCAAAT | TAGTACCATC | 1500 |
| | ATACTTCTTA | AATGAAGCGG | GATCCGCAAA | TGGTATCGTA | TCAATGATGG | GTGGTTTAGG | 1560 |
| | AGGATTCTTC | CCACCACTAG | TAATCACGTA | CGTAGCTAAT | TTAACAGGAT | CAAGTCATTT | .1620 |
| 50 | AGCATTTATT | TTCTTAGCGG | TATTnGGAnG | TATTGCATTA | TTTACCATGC | GTCATTTATA | 1680 |
| | CCAAAAAGAA | TATGGCTCAT | TGAAaAACGG | TTGATATGTA | ATACATGCCA | TTCATTTAGT | 1740 |
| | ********* | GCCTT= ATAT | CATGCGCAAT | ATTCGTAGCA | TGACATTAAG | GCTTTAGTAG | 1800 |